Number Restrictions on Transitive Roles in Description Logics with Nominals

Víctor Gutiérrez-Basulto
Cardiff University, UK
GutierrezBasultoV@cardiff.ac.uk

Yazmín Ibáñez-García
TU Wien, Austria
yazmin.garcia@tuwien.ac.at

Jean Christoph Jung
Universität Bremen, Germany
jeanjung@informatik.uni-bremen.de

Abstract

We study description logics (DLs) supporting number restrictions on transitive roles. We first take a look at $\mathcal{SOQ}$ and $\mathcal{SON}$ with binary and unary coding of numbers, and provide algorithms for the satisfiability problem and tight complexity bounds ranging from EXPTIME to NEXPTIME. We then show that by allowing for counting only up to one (functionality), inverse roles and role inclusions can be added without losing decidability. We finally investigate DLs of the *DL-Lite*-family, and show that, in the presence of role inclusions, the *core* fragment becomes undecidable.

1 Introduction

Description Logics (DLs) are a successful family of logic-based knowledge representation formalisms. The relevance of DLs comes from the fact that they are arguably the most popular language for the formulation of ontologies. For instance, they provide the logical basis of the web ontology language OWL 2, the medical ontology SNOMED CT, and the NCI thesaurus. One of the main reasons for the take-up of DLs is that, in general, they provide a good trade-off between expressivity and computational complexity. Unfortunately, in some cases this is not easy to ensure, e.g., the unrestricted interaction of *(qualified) number restrictions* and *transitive roles* tends to destroy this good balance; in many cases, leading to undecidability. On the other hand, support of these features is required, e.g., for ontological modeling in the biomedical domain (Rector and Rogers 2006; Kazakov, Sattler, and Zolin 2007; Stevens et al. 2007). For instance, in the classification of proteins (Wolstencroft et al. 2005), certain classes of proteins are defined in terms of their composition: *If a protein contains* at least n_1 X_1-*components* ... *and* at least n_k X_k *components, then it belongs to class* B. Moreover, these definitions require modeling of parthood, which is intended to be a *transitive relation*. Hence there is need of a clear understanding of the decidability frontier for DLs supporting these features.

With this in mind, in the last 15 years, the DL community has developed a vast amount of research on the complexity of reasoning in the presence of transitive roles and number restrictions, see, (Horrocks, Sattler, and Tobies 2000; Kazakov, Sattler, and Zolin 2007; Schröder and Pattinson 2008; Kaminski and Smolka 2010) and references therein. In particular, it has been shown that the extensions of $\mathcal{SN}$ ($\mathcal{ALC}$ enriched with transitive roles and *unqualified* counting) with role inclusions ($\mathcal{SHN}$) or inverse roles ($\mathcal{SIN}$) are undecidable (Horrocks, Sattler, and Tobies 2000; Kazakov, Sattler, and Zolin 2007). These negative results are, intuitively, explained by the interaction of these two constructors, i.e., by the possibility of counting over transitive roles. In order to regain decidability, different restrictions on their interaction have been proposed, e.g., to completely disallow number restrictions on transitive roles, or impose certain restrictions on the transitive roles occurring in role inclusions (Horrocks, Sattler, and Tobies 2000; Kazakov, Sattler, and Zolin 2007). On the positive side, it was shown that if role inclusions and inverse roles are not present, as in $\mathcal{SN}$, $\mathcal{SQ}$, $\mathcal{SOQ}$, decidability is then regained (Kazakov, Sattler, and Zolin 2007; Kaminski and Smolka 2010). However, no (elementary) complexity bounds are obtained from these results. Interestingly, if we have inverse roles or role inclusions, but only *functionality* (counting up to one) is allowed the panorama is less clear. The only known result is that satisfiability relative to $\mathcal{SIF}$-TBoxes is decidable in 2EXPTIME (Tendera 2005), but decidability of, e.g., $\mathcal{SHIF}, \mathcal{SHOIF}$ remains an open problem.

The main contribution of this paper is to establish a complete picture of the complexity of the problem of concept satisfiability relative to TBoxes in DLs supporting counting over transitive roles, by resolving the aforementioned open problems. Moreover, for all considered DLs including nominals, our upper bound results transfer to knowledge base satisfiability.

Our investigation starts (Section 3) with the DL $\mathcal{SOQ}$, allowing for qualified counting and nominals. As mentioned above, decidability was shown by Kaminski and Smolka [2010], and NEXPTIME-hardness is inherited from graded modal logic (Kazakov and Pratt-Hartmann 2009).[1] However, the exact computational complexity of $\mathcal{SOQ}$ was unknown. We close here this gap, by providing a NEXPTIME upper bound. To this aim, we use a two-step approach. First, we provide a decomposition of $\mathcal{SOQ}$ models, permitting us to 'independently reason' about the different (transitive) roles. In a second step, carefully adapting a tech-

[1] Graded modalities correspond to qualified number restrictions.

nique developed by Kazakov and Pratt-Hartmann [2009] in the context of graded modal logic, we show a small (that is, exponential) model property of each member of the decomposition, which lifts to $\mathcal{SOQ}$ and thus leads to the desired NEXPTIME upper bound.

As the next step (Section 4), we turn our attention to $\mathcal{SON}$, the restriction of $\mathcal{SOQ}$ to unqualified number restrictions. In particular, we are interested in understanding the impact of the coding of numbers on the computational complexity. We first show that with unary coding, satisfiability in $\mathcal{SON}$ is EXPTIME-complete, and therefore easier than in $\mathcal{SOQ}$. We devise a type-elimination procedure that exploits the unary coding by the observation that certain witnesses are of only polynomial size, and can thus be all enumerated. We then show that with binary coding, the complexity of satisfiability jumps to NEXPTIME-complete. In fact, the lower bound holds already for concept satisfiability of $\mathcal{SN}$ concepts over a single transitive role (no TBox). It is interesting to note that when only *non-transitive* roles are allowed in number restrictions the coding has no impact on the computational complexity, that is, regardless of the coding of numbers, satisfiability in $\mathcal{SON}$ and $\mathcal{SOQ}$ is EXPTIME-complete (Calvanese, Eiter, and Ortiz 2009).

We then take a look (Section 5) at the case when only functionality is allowed. We show that in this case inverse roles, role inclusions and nominals can be added without losing decidability. In particular, we show that satisfiability in $\mathcal{SHIF}$ and $\mathcal{SHOIF}$ is EXPTIME- and NEXPTIME-complete, respectively, and hence not harder than when one cannot impose number restrictions on transitive roles.

Lightweight DLs of the *DL-Lite* family allowing for number restrictions on transitive roles have not been considered yet; indeed, only counting over non-transitive roles has been studied in *DL-Lite* (Artale et al. 2009). In the last part of the paper (Section 6), we initiate this study. In particular, we complement known undecidability results by considering a light sub-Boolean DL with unqualified existential restrictions and show that the *core* fragment of *DL-Lite*, with role inclusions, allowing for number restrictions on transitive roles is undecidable.

Missing proofs are available at www.informatik.uni-bremen.de/tdki/research/papers/GIJ17.pdf .

2 Preliminaries

Syntax. We introduce the DL $\mathcal{SHOIQ}$ (Hollunder and Baader 1991), which extends the classical DL $\mathcal{ALC}$ with transitivity declarations on roles ($\mathcal{S}$), role inclusion axioms ($\mathcal{H}$), nominals ($\mathcal{O}$), inverses ($\mathcal{I}$), and qualified number restrictions ($\mathcal{Q}$). We consider a vocabulary consisting of countably infinite disjoint sets of *concept names* $\mathsf{N_C}$, *role names* $\mathsf{N_R}$ and *individual names* $\mathsf{N_I}$. The syntax of $\mathcal{SHOIQ}$*-concepts* C, D is given by the following grammar:

$$C, D ::= A \mid \neg C \mid C \sqcap D \mid \{o\} \mid \exists r.C \mid (\sim n\; r\; C)$$

where $A \in \mathsf{N_C}$, $o \in \mathsf{N_I}$, $r \in \{s, s^- \mid s \in \mathsf{N_R}\}$ is a *role*, $\sim$ is a comparison operator $\leq$ or $\geq$, and n is a number (given in binary, unless stated otherwise). Roles of the form r^- are called *inverse roles*, concepts of the form $\{o\}$, $\exists r.C$, $(\leq n\; r\; C)$, $(\geq n\; r\; C)$ are called, respectively, *nominals*, *existential restrictions*, *at most-restrictions* and *at least-restrictions*. We identify r^- with $s \in \mathsf{N_R}$ if $r = s^-$, and use standard abbreviations $\top$, $\bot$, $C \sqcup D$, $\forall r.C$, and $C \rightarrow D$.

A $\mathcal{SHOIQ}$*-TBox (ontology)* $\mathcal{T}$ is a finite set of *concept inclusions (CIs)* $C \sqsubseteq D$, *transitivity declarations* $\mathsf{Tra}(r)$ and *role inclusions (RIs)* $r \sqsubseteq s$, where C, D are $\mathcal{SHOIQ}$-concepts and r, s roles. We use $\mathsf{CN}(\mathcal{T})$, $\mathsf{Rol}(\mathcal{T})$ and $\mathsf{Nom}(\mathcal{T})$ to denote, respectively, the set of *all concept names, roles and nominals occurring in* $\mathcal{T}$. Wlog. we assume that if $\mathsf{Tra}(r) \in \mathcal{T}$ then $\mathsf{Tra}(r^-) \in \mathcal{T}$. Indeed, by the semantics, if a role is transitive, so is its inverse.

Semantics. As usual, the semantics is defined in terms of interpretations. An *interpretation* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ consists of a non-empty *domain* $\Delta^{\mathcal{I}}$ and an *interpretation function* $\cdot^{\mathcal{I}}$ mapping concept names to subsets of the domain and role names to binary relations over the domain. We define, mutually recursive, the set $r_{\mathcal{I}}(d, C) = \{e \in C^{\mathcal{I}} \mid (d, e) \in r^{\mathcal{I}}\}$ of r-successors of d satisfying C, and the interpretation of complex concepts $C^{\mathcal{I}}$ by taking

$$\begin{aligned}
(r^-)^{\mathcal{I}} &= \{(e, d) \mid (d, e) \in r^{\mathcal{I}}\};\\
(\neg C)^{\mathcal{I}} &= \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}};\\
(C \sqcap D)^{\mathcal{I}} &= C^{\mathcal{I}} \cap D^{\mathcal{I}};\\
\{o\}^{\mathcal{I}} &= \{o^{\mathcal{I}}\};\\
(\exists r.C)^{\mathcal{I}} &= \{d \in \Delta^{\mathcal{I}} \mid \exists e \in C^{\mathcal{I}} \text{ with } (d, e) \in r^{\mathcal{I}}\};\\
(\sim n\; r\; C)^{\mathcal{I}} &= \{d \in \Delta^{\mathcal{I}} \mid |r_{\mathcal{I}}(d, C)| \sim n\}.
\end{aligned}$$

The satisfaction relation $\models$ is defined standardly:

$$\mathcal{I} \models C \sqsubseteq D \text{ iff } C^{\mathcal{I}} \subseteq D^{\mathcal{I}};$$
$$\mathcal{I} \models r \sqsubseteq s \text{ iff } r^{\mathcal{I}} \subseteq s^{\mathcal{I}};$$
$$\mathcal{I} \models \mathsf{Tra}(r) \text{ iff } r^{\mathcal{I}} \text{ is transitive.}$$

An interpretation $\mathcal{I}$ is a *model* of a TBox $\mathcal{T}$, denoted $\mathcal{I} \models \mathcal{T}$, if $\mathcal{I} \models \alpha$ for all $\alpha \in \mathcal{T}$. A concept C is *satisfiable relative to a TBox* $\mathcal{T}$ if there is a model $\mathcal{I}$ of $\mathcal{T}$ such that $C^{\mathcal{I}} \neq \emptyset$.

Reasoning Problem. We are interested in the problem of *concept satisfiability*, that is, given a TBox $\mathcal{T}$ and a concept C, we want to determine whether C is satisfiable relative to $\mathcal{T}$. We restrict our attention to the case when $C = A \in \mathsf{N_C}$ because C is satisfiable relative to $\mathcal{T}$ iff A_C is satisfiable relative to $\mathcal{T} \cup \{A_C \sqsubseteq C\}$ for any fresh concept name A_C.

Please note that in the presence of nominals our upper bounds transfer to the problem of *knowledge base satisfiability*; indeed, so-called *ABox assertions* can be internalized in the TBox using nominals (Baader et al. 2003).

Fragments. We consider the following fragments:

- $\mathcal{SOQ}$ is obtained from $\mathcal{SHOIQ}$ by disallowing role inclusions and inverse roles.
- $\mathcal{SON}$ is obtained from $\mathcal{SOQ}$ by supporting only *unqualified* number restrictions (indicated by letter $\mathcal{N}$) of the form $(\sim n\; r\; \top)$, which we usually abbreviate as $(\sim n\; r)$.
- $\mathcal{SHOIF}$ is obtained from $\mathcal{SHOIQ}$ by supporting only *local functionality* constraints (indicated by letter $\mathcal{F}$) of the form $(\leq 1\; r)$.

3 $\mathcal{SOQ}$

We start by devising an algorithm for concept satisfiability relative to $\mathcal{SOQ}$-TBoxes, yielding a tight NEXPTIME upper bound. The matching lower bound follows from the fact that satisfiability in the graded modal logic over transitive frames, **GrK4**, is NEXPTIME-complete (Kazakov and Pratt-Hartmann 2009).

We assume that the input TBox is in the following normal form. Let $\mathcal{C}_{\mathsf{Bool}}$ be the set of $\mathcal{SOQ}$-concept descriptions that are obtained without using the constructors $(\sim n\ r\ C)$ and $\exists r.C$ (we treat $\exists r.C$ as $(\geq 1\ r\ C)$). We say that a TBox is in normal form if all concept inclusions are of the shape

$$C \sqsubseteq D \quad \text{or} \quad C \sqsubseteq (\sim n\, r\, D),$$

for $C, D \in \mathcal{C}_{\mathsf{Bool}}$. We show that, by introducing fresh concept names, every TBox can be transformed in polynomial time into a satisfiability-equivalent TBox in normal form.

In order to obtain the desired NEXPTIME upper bound it clearly suffices to show a *small model property*, that is, whenever A is satisfiable relative to $\mathcal{T}$, then there is a model of exponential size, since we can then simply "guess" the model. To this end, we will first characterize concept satisfiability in terms of the existence of a *quasimodel*, which is a decomposition of a model of $\mathcal{SOQ}$-TBoxes into components that interpret only a single role name. This is in line with viewing $\mathcal{SOQ}$ as a fusion logic. Note that decompositions of fusion logics have been studied (Baader et al. 2002), but so far nominals were not considered. Nominals impose the additional difficulty that the models of a $\mathcal{SOQ}$-TBox are not closed under union.

Let $\mathcal{T}$ be the input TBox. For $r \in \mathsf{N_R}$, we define $\mathcal{T}_r := \mathcal{T} \setminus \{C \sqsubseteq (\sim n\ r'\ D) \mid r' \neq r\}$. Intuitively, $\mathcal{T}_r$ reflects $\mathcal{T}$ on the single role r. Given two interpretations $\mathcal{I}, \mathcal{J}$, we say that $d \in \Delta^{\mathcal{I}}, d' \in \Delta^{\mathcal{J}}$ are *Boolean equivalent* iff for all $C \in \mathcal{C}_{\mathsf{Bool}}$ we have $d \in C^{\mathcal{I}}$ iff $d' \in C^{\mathcal{J}}$.

We are now in a position to define the intended decomposition. A *quasimodel for* $\mathcal{T}$ is a finite collection of interpretations $\mathfrak{Q} = \{\mathcal{I}_r \mid r \in \mathsf{Rol}(\mathcal{T})\}$ such that the following two conditions are satisfied:

(**qm1**) $\mathcal{I}_r \models \mathcal{T}_r$ for each role name $r \in \mathsf{Rol}(\mathcal{T})$;

(**qm2**) for all role names r, s and $d \in \Delta^{\mathcal{I}_r}$ there exists a $d' \in \Delta^{\mathcal{I}_s}$ such that d and d' are Boolean equivalent.

Intuitively, (**qm1**) captures the TBox relative to a single role name r, and (**qm2**) ensures that the components $\mathcal{I}_r$ can be combined into one model. A *quasimodel for* A *and* $\mathcal{T}$ is a quasimodel $\mathfrak{Q}$ for $\mathcal{T}$ such that $A^{\mathcal{I}_r} \neq \emptyset$ for some (equivalently: every) interpretation $\mathcal{I}_r \in \mathfrak{Q}$. Note that, by (**qm1**), we can assume that each $\mathcal{I}_r$ interprets only role r as (possibly) non-empty. Thus, quasimodels provide a suitable decomposition of models. The *size* of a quasimodel is the sum of the domain sizes of all interpretations in the quasimodel.

The following lemma provides the characterization of satisfiability and additionally relates the size of the quasimodel to the size of a model.

Lemma 1. *A is satisfiable relative to $\mathcal{T}$ iff there is a quasimodel for A and $\mathcal{T}$. Moreover, if there is a quasimodel of size κ for A and $\mathcal{T}$, there is a model of size $\leq \kappa$ for A and $\mathcal{T}$.*

It remains to restrict the sizes of quasimodels.

Lemma 2. *If there is a quasimodel for A and $\mathcal{T}$, there is a quasimodel for A and $\mathcal{T}$ of exponential size.*

We give some intuitions on the proof here. Let $\mathfrak{Q}$ be a quasimodel for A and $\mathcal{T}$ and $\mathcal{I}_r \in \mathfrak{Q}$. If r is a non-transitive role, it has been already shown that $\mathcal{I}_r$ can be replaced by an exponentially sized interpretation $\mathcal{I}'_r$, preserving (**qm2**) (Lutz et al. 2005, Corollary 4.3). Therefore, we concentrate on the case when r is a transitive role, that is, $\mathsf{Tra}(r) \in \mathcal{T}$.

First observe that we can assume that $\mathcal{I}_r$ has at most exponentially many connected components, more precisely, $2^{|X|}$, where $X = \mathsf{CN}(\mathcal{T}) \cup \mathsf{Nom}(\mathcal{T})$. To see this, fix for every subset $Y \subseteq X$ a domain element d_Y such that $Y = \{C \in X \mid d_Y \in C^{\mathcal{I}_r}\}$, if such an element exists. It should be clear that the restriction $\mathcal{I}'$ of $\mathcal{I}_r$ to domain

$$\Delta^{\mathcal{I}'} = \{d \in \Delta^{\mathcal{I}} \mid \exists d_Y : (d_Y, d) \in (r^{\mathcal{I}})^*\}$$

still satisfies (**qm1**) and (**qm2**), and additionally has at most $2^{|X|}$ connected components (each rooted at some d_Y).

Finally, to show Lemma 2, we carefully adapt a technique by Kazakov and Pratt-Hartmann [2009] showing the finite model property of **GrK4**, to prove that every connected component of $\mathcal{I}_r$ can be assumed to be of exponential size. Crucially, we have to take care that, when domain elements are removed, we keep the witnesses d_Y for Condition (**qm2**). See the appendix for a full proof.

Lemma 1 and 2 yield the small (that is, exponential) model property for $\mathcal{SOQ}$ which, as argued above implies:

Theorem 1. *Concept satisfiability relative to $\mathcal{SOQ}$-TBoxes is* NEXPTIME*-complete.*

4 $\mathcal{SON}$

In this section, we study the complexity of concept satisfiability relative to $\mathcal{SON}$-TBoxes, with both unary and binary coding of numbers. Note that for $\mathcal{SOQ}$ the coding of numbers does not make any difference on the computational complexity because the NEXPTIME-hardness proof for satisfiability in **GrK4** only uses numbers that are at most 1.

We first show that with unary coding, concept satisfiability relative to $\mathcal{SON}$-TBoxes is EXPTIME-complete, and thus easier than in $\mathcal{SOQ}$. We then show that with binary coding, the complexity of concept satisfiability relative to $\mathcal{SON}$-TBoxes coincides with that relative to $\mathcal{SOQ}$-TBoxes. In particual, we show that the latter holds already for $\mathcal{SN}$.

4.1 Unary Coding of Numbers

We focus on providing an EXPTIME algorithm for concept satisfiability relative to $\mathcal{SON}$-TBoxes with unary coding of numbers. The lower bound is inherited from $\mathcal{ALC}$.

We proceed in two steps. First, we give a characterization of concept satisfiability, independent of the coding of numbers. This characterization is then the basis for a type elimination procedure which runs in exponential time, given the unary coding. The main challenge lies in the interplay between nominals and transitive roles. In fact, the algorithm is not purely type-based, but needs to make explicit what we call the *nominal core* of an interpretation, which is the part of the model 'close' to the nominals.

Denote with $\mathsf{cl}(\mathcal{T})$ the set of all sub-concepts appearing in $\mathcal{T}$, closed under single negations. A *type for* $\mathcal{T}$ is a set $t \subseteq \mathsf{cl}(\mathcal{T})$ satisfying

– $D \in t$ iff $\neg D \notin t$ for all $\neg D \in \mathsf{cl}(\mathcal{T})$;

– $D \sqcap E \in t$ iff $\{D, E\} \subseteq t$, for all $D \sqcap E \in \mathsf{cl}(\mathcal{T})$;

– $C \in t$ implies $D \in t$, for all $C \sqsubseteq D \in \mathcal{T}$.

Let $\mathsf{tp}(\mathcal{T})$ be the set of all types for $\mathcal{T}$. Two types $t, t' \in \mathsf{tp}(\mathcal{T})$ are called *r-compatible*, written $t \rightsquigarrow_r t'$, if

– $\{\neg D \mid \neg\exists r.D \in t\} \subseteq t'$, in case r is non-transitive, and

– $\{\neg D, \neg\exists r.D \mid \neg\exists r.D \in t\} \subseteq t'$, in case r is transitive.

Fix a role r and a type t, and let ℓ be maximal with $(\geq \ell\ r) \in t$, and u be minimal with $(\leq u\ r) \in t$.[2] Then, t is called *r-realizable in* $T \subseteq \mathsf{tp}(\mathcal{T})$ if $\ell \leq u$ and there are $k \leq u$ types $t_1, \dots, t_k \in T$ with $t \rightsquigarrow_r t_i$, for all i, such that:

1. for each $\exists r.C \in t$, there is some i with $C \in t_i$;
2. if $k < \ell$, there is an i with $\{o\} \notin t_i$ for all $\{o\} \in \mathsf{cl}(\mathcal{T})$.

Item 1 states the known realizability condition for $\mathcal{ALC}$. Item 2 captures the interplay of nominals and *at-least* restrictions; in particular, if $k < \ell$, we need one type that can be repeated as a successor, which cannot be a nominal type.

For transitive roles in combination with *at-most* restrictions $(\leq n\ r)$ and nominals, we need to make explicit how the at-most restrictions in a type are realized. In order to formalize this, denote with $\mathsf{tp}_\mathcal{I}(d)$ the type $\{C \in \mathsf{cl}(\mathcal{T}) \mid d \in C^\mathcal{I}\}$ of d in an interpretation $\mathcal{I}$, and say that $t =_r t'$ if $C \in t$ iff $C \in t'$ for all concepts $C \in \mathsf{cl}(\mathcal{T})$ *not* of the form $(\sim n\ s)$, $\neg(\sim n\ s)$, $\exists s.A$, $\neg\exists s.A$, for $s \neq r$. Moreover, given a role name r and sets $T' \subseteq T \subseteq \mathsf{tp}(\mathcal{T})$, we say that an interpretation $\mathcal{I}$ is a *$\leq$-witness for* (r, T, T') if

(i) for each $d \in \Delta^\mathcal{I}$, there is $t \in T$ with $\mathsf{tp}_\mathcal{I}(d) =_r t$, and

(ii) for every $t \in T'$ such that $(\leq n\ r) \in t$, there is some d with $\mathsf{tp}_\mathcal{I}(d) =_r t$.

Intuitively, $\mathcal{I}$ realizes (relative to r) only types from T, but at least those in T'. Using the notion of $\leq$-witness we give the following characterization of concept satisfiability, which also provides the starting point of our decision procedure.

Lemma 3. *A is satisfiable relative to $\mathcal{T}$ iff there is a set $T \subseteq \mathsf{tp}(\mathcal{T})$ with $A \in t$ for some $t \in T$ such that:*

(E1) *for any $\{o\}$ in $\mathcal{T}$ there is exactly one $t \in T$ with $\{o\} \in t$;*

(E2) *every $t \in T$ is r-realizable in T, for each role r;*

(E3) *there is a $\leq$-witness for (r, T, T), for any transitive r.*

Conditions (**E1**) and (**E2**) are straightforward; Condition (**E3**) requires, for each transitive role r, an interpretation realizing all types with an at-most restriction. Though condition (**E3**) is intuitive, it does not lend itself for implementation yet because $\leq$-witnesses are exponentially big in general.

[2] By convention, $\ell = 0$ and $u = \infty$, respectively, if no such concepts are in t.

As the next step, we analyze the $\leq$-witnesses and give an equivalent condition (**E3'**). Fix a role name r. The *nominal core of an interpretation $\mathcal{I}$ wrt. r*, written $\mathsf{core}_r(\mathcal{I})$, is obtained from $\mathcal{I}$ by restricting the domain to

$$\{o^\mathcal{I} \mid \{o\} \in \mathsf{cl}(\mathcal{T})\} \cup$$
$$\{d \mid (o^\mathcal{I}, d) \in r^\mathcal{I}, o^\mathcal{I} \in (\leq m\ r)^\mathcal{I}, (\leq m\ r) \in \mathsf{cl}(\mathcal{T})\}$$

We prove that Lemma 3 remains correct if we use the following instead of (**E3**):

(**E3'**) for each transitive r, there is an interpretation $\mathcal{I}_r$ such that, for each $t \in T$ with $(\leq n\ r) \in t$, there is a $\leq$-witness $\mathcal{I}_{rt}$ for $(r, T, \{t\})$ with $\mathsf{core}_r(\mathcal{I}_{rt}) = \mathcal{I}_r$.

Before we give the algorithm, observe that there are only exponentially many *maximal* sets $T \subseteq \mathsf{tp}(\mathcal{T})$ satisfying (**E1**), that is, there is no set T' satisfying (**E1**) and $T \subsetneq T' \subseteq \mathsf{cl}(\mathcal{T})$. Moreover, it is crucial to observe that the (domain) size of the nominal core of a $\leq$-witness (and in fact of any interpretation) is *polynomial in the size of* $\mathcal{T}$; more precisely, its size is bounded by $\ell_1\ell_2$, where ℓ_1 is the number of nominals in $\mathcal{T}$, and ℓ_2 is the largest number in $\mathcal{T}$; so the $\mathcal{I}_r$ in (**E3'**) has only polynomial size. Finally, given such a (polynomial) $\mathcal{I}_r$ and some t with $(\leq n\ r) \in t$, we can check in exponential time the existence of $\mathcal{I}_{rt}$, because we can just try all possible extensions of $\mathcal{I}_r$ with n elements.

These arguments show that the following procedure runs in exponential time. For each maximal set $T \subseteq \mathsf{cl}(\mathcal{T})$ satisfying (**E1**) and each possible combination of nominal cores (one for each role name), exhaustively remove types from T if they do not satisfy (**E2**) or (**E3'**). Accept if, in this way, a set $\hat{T}$ is found which satisfies all conditions and there is $t \in \hat{T}$ with $A \in t$. Overall, this shows:

Theorem 2. *Concept satisfiability relative to $\mathcal{SON}$-TBoxes with unary coding of numbers is* EXPTIME*-complete.*

4.2 Binary Coding of Numbers

Now, we show that with binary coding, concept satisfiability relative to $\mathcal{SON}$-TBoxes becomes NEXPTIME-hard. The matching upper bound follows from Theorem 1 above. Note that the lower bound does not follow from (the proof of) NEXPTIME-hardness of satisfiability in **GrK4** because that relies on qualified number restrictions.

The NEXPTIME-hardness proof is by reduction of the problem of tiling a torus of exponential size (van Emde Boas 1997). For the reduction to work nominals are not required, that is, the lower bound already holds for $\mathcal{SN}$. Intuitively, in the reduction, we cope with the lack of qualified number restrictions by exploiting the fact that 'big numbers' can be used due to the binary coding.

We concentrate here on the most interesting part, the construction of an $\mathcal{SN}$-TBox $\mathcal{T}_{\mathsf{tor}}$ and a concept L_0 whose satisfiability characterize the $2^n \times 2^n$-torus. To this aim, we use the following signature:

- concept names $X_0, \dots X_{n-1}$, $Y_0, \dots, Y_{n-1}$ that serve to encode the (x, y) coordinates in the torus
- concept names $L_0, \dots, L_{2n}$ that mark the levels of a binary tree,

- a transitive role r.

We start by enforcing that certain models of $\mathcal{T}_{\mathsf{tor}}$ contain a complete binary tree, called *torus-tree*, with $2n$ levels, where the 2^{2n} leaves of the torus-tree will represent the $2^n \times 2^n$ points in the torus. To this end, we include the transitivity statement $\mathsf{Tra}(r)$ and the following CIs, for $0 \leq i < n$ in $\mathcal{T}_{\mathsf{tor}}$:

$$L_i \sqsubseteq \exists r.(X_i \sqcap L_{i+1}) \sqcap \exists r.(\neg X_i \sqcap L_{i+1})$$
$$L_{n+i} \sqsubseteq \exists r.(Y_i \sqcap L_{n+i+1}) \sqcap \exists r.(\neg Y_i \sqcap L_{n+i+1}).$$

Moreover, we force the levels to be disjoint by adding the CI

$$L_i \sqsubseteq \neg L_j, \qquad \text{for } 0 \leq i < j \leq 2n.$$

We now propagate the concepts X_i and Y_i down to level L_{2n} to encode numbers between 0 and $2^n - 1$ at the leaves of the torus-tree. Since r is a transitive role, the following concept inclusions, for all i with $0 \leq i < n$ suffice:

$$L_{i+1} \sqcap X_i \sqsubseteq \forall r.(L_j \to X_i) \qquad \text{for } i < j \leq 2n,$$
$$L_{i+1} \sqcap \neg X_i \sqsubseteq \forall r.(L_j \to \neg X_i) \qquad \text{for } i < j \leq 2n,$$
$$L_{n+i+1} \sqcap Y_i \sqsubseteq \forall r.(L_j \to Y_i) \qquad \text{for } n+i < j \leq 2n,$$
$$L_{n+i+1} \sqcap \neg Y_i \sqsubseteq \forall r.(L_j \to \neg Y_i) \qquad \text{for } n+i < j \leq 2n.$$

Next, we introduce some required notation. Fix an interpretation $\mathcal{I}$. For each element $d \in \Delta^{\mathcal{I}}$, we define $\mathsf{pos}(d)$ as the pair of integers

$$(\mathsf{xpos}(d), \mathsf{ypos}(d)) = (\Sigma_0^{n-1} x_i \cdot 2^i, \Sigma_0^{n-1} y_i \cdot 2^i),$$

where

$$x_i = \begin{cases} 0 & \text{if } d \notin X_i^{\mathcal{I}}, \\ 1 & \text{otherwise;} \end{cases} \qquad y_i = \begin{cases} 0 & \text{if } d \notin Y_i^{\mathcal{I}}, \\ 1 & \text{otherwise.} \end{cases}$$

It should be clear that in any model $\mathcal{I}$ of L_0 and the CIs defined so far, there are 2^{2n} elements which satisfy L_{2n}; even more, for each pair of values $0 \leq i, j < 2^n$, there is an element $d_{ij} \in L_{2n}^{\mathcal{I}}$ such that $\mathsf{pos}(d_{ij}) = (i, j)$. However, the elements d_{ij} are not necessarily connected in a particularly useful way; thus, we now relate elements at level $2n$ to their horizontal and vertical neighbors.

To this aim, we will use *glueing points*. More precisely, for every $d, d' \in L_{2n}^{\mathcal{I}}$ with $\mathsf{pos}(d) = (x, y)$ and $\mathsf{pos}(d') = (x \oplus_{2^n} 1, y)$,[3] we enforce an element $g \in H^{\mathcal{I}}$ such that $(d, g) \in r^{\mathcal{I}}$ and $(d', g) \in r^{\mathcal{I}}$ and $\mathsf{pos}(g) = (x \oplus_{2^n} 1, y)$, and similar for the y-coordinate. This is illustrated in Figure 1, where glueing points are depicted as $\circ$ and labelled with H and V for horizontal and vertical, respectively.

To facilitate this task, we define the following concepts, for $0 \leq i \leq n-1$ and $i < j \leq n-1$:

$$X_i^* \equiv \neg X_i \sqcap \bigsqcap_{0 \leq k \leq i-1} X_k; \qquad X_i^+ \equiv X_i \sqcap \bigsqcap_{0 \leq k \leq i-1} \neg X_k;$$
$$X_n^* \equiv \bigsqcap_{0 \leq k \leq n-1} X_k; \qquad X_n^+ \equiv \bigsqcap_{0 \leq k \leq n-1} \neg X_k;$$
$$X_i^{\to} \sqsubseteq (X_j \to \forall r.X_j) \sqcap (\neg X_j \to \forall r.\neg X_j));$$

[3] $\oplus_k$ denotes the addition modulo k.

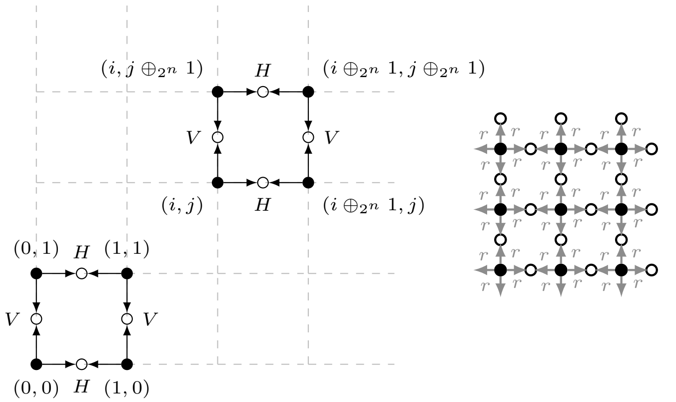

Figure 1: Glueing the points in the torus

and analogous concepts Y_i^*, Y_i^+, and $Y_i^{\to}$. Observe that for every interpretation $\mathcal{I}$ and $d \in \Delta^{\mathcal{I}}$, there is a *single* i such that $d \in (X_i^+)^{\mathcal{I}}$, and similarly for X_i^*. Moreover, for every $d, e \in \Delta^{\mathcal{I}}$ with $\mathsf{xpos}(e) = \mathsf{xpos}(d) \oplus_{2^n} 1$, we have that, if $d \in (X_i^*)^{\mathcal{I}}$, then $e \in (X_i^+)^{\mathcal{I}}$ and $d \in X_j^{\mathcal{I}}$ iff $e \in X_j^{\mathcal{I}}$, for all $i < j \leq n$; and similarly for ypos. With this in mind, we enforce for every element d in level $2n$ with $\mathsf{pos}(d) = (x, y)$ four r-successors $d_h^+, d_h^=, d_v^+, d_v^=$ such that

- $d_h^+ \in H^{\mathcal{I}}$, $\mathsf{xpos}(d_h^+) = x \oplus_{2^n} 1$, $\mathsf{ypos}(d_h^+) = y$,
- $d_h^= \in H^{\mathcal{I}}$, $\mathsf{xpos}(d_h^=) = x$, $\mathsf{ypos}(d_h^=) = y$,
- $d_v^+ \in V^{\mathcal{I}}$, $\mathsf{xpos}(d_v^+) = x$, $\mathsf{ypos}(d_v^+) = y \oplus_{2^n} 1$, and
- $d_v^= \in V^{\mathcal{I}}$, $\mathsf{xpos}(d_v^=) = x$, $\mathsf{ypos}(d_v^=) = y$,

using the following concept inclusions, for $0 \leq i \leq n$ and $0 \leq j \leq n-1$:

$$L_{2n} \sqcap X_i^* \sqsubseteq X_i^{\to} \sqcap \exists r.(H \sqcap X_i^+) \sqcap \exists r.(H \sqcap X_i^*)$$
$$L_{2n} \sqcap Y_i^* \sqsubseteq Y_i^{\to} \sqcap \exists r.(V \sqcap Y_i^+) \sqcap \exists r.(V \sqcap Y_i^*)$$
$$L_{2n} \sqcap X_j \sqsubseteq \forall r.(V \to X_j)$$
$$L_{2n} \sqcap \neg X_j \sqsubseteq \forall r.(V \to \neg X_j)$$
$$L_{2n} \sqcap Y_j \sqsubseteq \forall r.(H \to Y_j)$$
$$L_{2n} \sqcap \neg Y_j \sqsubseteq \forall r.(H \to \neg Y_j)$$

Moreover, we make sure that the glueing points are fresh, and that horizontal and vertical are disjoint by adding:

$$H \sqsubseteq \neg V \quad \text{and} \quad H \sqcup V \sqsubseteq \neg L_i, \text{for } 0 \leq i \leq 2n.$$

It remains to *identify* the introduced glueing points as indicated in Figure 1. For this, we use the (unqualified) number restrictions. In particular, we add the concept inclusion

$$L_0 \sqsubseteq (\leq k\ r),$$

with $k = (2^{2n+1} - 2) + 2^{2n}$. To justify the choice of k, note that, without the glueing points, the intended model of L_0 has 2^i elements in every level i, that is, $2^{2n+1} - 1$ elements overall, and hence L_0 has $2^{2n+1} - 2$ successors. As we want to have a single glueing point for every (i, j) with $0 \leq i, j < n$, we need to restrict the number of glueing points to 2^{2n}.

This finishes the definition of $\mathcal{T}_{\mathsf{tor}}$. It is formally shown in the appendix that $\mathcal{T}_{\mathsf{tor}}$ properly defines the $2^n \times 2^n$-torus.

Having this, it is standard to reduce the tiling problem on the torus, by using the elements in level $2n$ as the tiles and the glueing points to communicate between neighboring tiles.

Theorem 3. *Concept satisfiability relative to $\mathcal{SON}$-TBoxes with binary coding of numbers is* NEXPTIME*-complete.*

Note that, in presence of a single transitive role r, we can always rewrite the TBox $\mathcal{T}$ as a concept $C_\mathcal{T}$: add a conjunct $(C \rightarrow D) \sqcap \forall r.(C \rightarrow D)$ for each $C \sqsubseteq D \in \mathcal{T}$; hence:

Corollary 1. *Given a transitive r, satisfiability of $\mathcal{SON}$-concepts with binary coding is* NEXPTIME*-complete.*

5 The Case of Functionality

We next study $\mathcal{SHOIF}$ which allows for both inverse roles and role inclusions. Recall that these features lead to undecidability already with unqualified counting with numbers greater than 1 (Horrocks, Sattler, and Tobies 2000; Kazakov, Sattler, and Zolin 2007). However, it was open whether decidability could be attained by sticking to functionality. We answer positively this question, by reducing satisfiability in $\mathcal{SHOIF}$ to satisfiability in $\mathcal{ALCHOIF}_{\mathsf{sf}}$, the extension of $\mathcal{ALCHOIF}$ with *local reflexivity* concepts $\exists r.\mathsf{self}$, whose semantics is given by

$$(\exists r.\mathsf{self})^\mathcal{I} = \{d \in \Delta^\mathcal{I} \mid (d,d) \in r^\mathcal{I}\}.$$

Let $\mathcal{T}$ be a $\mathcal{SHOIF}$-TBox and recall the notation $\mathsf{cl}(\mathcal{T})$ from Section 4.1. We obtain the TBox $\mathcal{T}'$ from $\mathcal{T}$ by taking $\mathcal{T}' = (\mathcal{T} \setminus \{\mathsf{Tra}(r) \in \mathcal{T}\}) \cup \mathcal{T}''$, where $\mathcal{T}''$ is the set of the following CIs, for every $\mathsf{Tra}(r) \in \mathcal{T}$ and $C \in \mathsf{cl}(\mathcal{T})$:

$$\forall r.C \sqsubseteq \forall r.(\forall r.C),$$
$$(\leq 1\ r) \sqsubseteq \forall r.(\forall r.\bot \sqcup \exists r.\mathsf{self}).$$

Concept inclusions of the first type have been used to mimic the behavior of transitive roles for example in (Tobies 2001); note also the similarity to the axiom $\Box p \rightarrow \Box\Box p$ characterizing transitive frames in modal logic (Chagrov and Zakharyaschev 1997). Concept inclusions of the second type capture the interplay of transitivity and (local) functionality: if two elements d and e are r-connected, for some functional transitive r, e cannot have an r-successor other than itself.

The correctness of this reduction is established in the appendix. It is easy to see that $\mathcal{T}'$ is an $\mathcal{ALCHOIF}_{\mathsf{sf}}$ TBox and that it can computed in polynomial time. Moreover, the reduction also works for $\mathcal{SHIF}$ yielding an $\mathcal{ALCHIQ}_{\mathsf{sf}}$-TBox. It is known that concept satisfiability in $\mathcal{ALCHIQ}_{\mathsf{sf}}$ and $\mathcal{ALCHOIQ}_{\mathsf{sf}}$ can be checked in EXPTIME (Calvanese, Eiter, and Ortiz 2009) and NEXPTIME (Motik, Shearer, and Horrocks 2009), respectively. Matching lower bounds are inherited from $\mathcal{ALC}$ and $\mathcal{ALCFIO}$, respectively (Baader et al. 2003; Lutz 2004).

Theorem 4. *Concept satisfiability is* NEXPTIME*-complete for $\mathcal{SHOIF}$-TBoxes and* EXPTIME*-complete for $\mathcal{SHIF}$.*

6 A Look at *DL-Lite*

We next show that supporting number restrictions on transitive roles in *DL-Lite*$^{\mathcal{HN}}_{core}$ (Artale et al. 2009) leads to undecidability. This result strengthens the known undecidability result for $\mathcal{SHIN}$ in the sense that *DL-Lite*$^{\mathcal{HN}}_{core}$ is a very weak sub-Boolean logic without qualified existential restrictions. *DL-Lite*$^{\mathcal{SHN}}_{core}$*-concepts* C are defined as

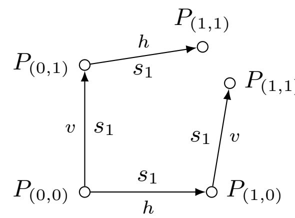

Figure 2: Grid Square

$$C ::= \bot \mid A \mid (\sim n\ r),$$

where $A \in \mathsf{N_C}$, r is a role and $\sim$ is an arbitrary comparison. *DL-Lite*$^{\mathcal{SHN}}_{core}$*-TBoxes* are defined as in Section 2, but CIs can only take the form: $C \sqsubseteq D$ or $C \sqcap D \sqsubseteq \bot$ with C, D *DL-Lite*$^{\mathcal{SHN}}_{core}$-concepts.

The undecidability proof (cf. appendix) is by reduction of the halting problem of deterministic Turing machines. In the proof, RIs and counting over transitive roles are key for the construction of squares of a grid, and for ensuring that such grid is infinite. For instance, in Figure 2, if we declare (i) the transitive role s_1 as super-role of h and v, and (ii) that each element has at-most 3 s_1-successors, then the two elements in $P_{(1,1)}$ are forced to be the same. Roughly, we then arrange a sequence of configurations (a computation) as a 'two-dimensional' grid of domain elements.

Theorem 5. *Concept satisfiability relative to DL-Lite*$^{\mathcal{SHN}}_{core}$*-TBoxes is undecidable.*

Decidability is regained again for functionality. In particular, Theorem 4 yields a (tight) EXPTIME upper bound for *DL-Lite*$^{\mathcal{SHF}}_{bool}$, which is the fragment of $\mathcal{SHIF}$ allowing only for *unqualified* existential restrictions. Note that the upper bound holds for local functionality; indeed, in *DL-Lite* functionality is normally meant to be *global*, which is weaker than the local one. The lower bound is inherited from *DL-Lite*$^{\mathcal{HF}}_{bool}$ (Artale et al. 2009). We thus obtain:

Theorem 6. *Concept satisfiability relative to DL-Lite*$^{\mathcal{SHF}}_{bool}$*-TBoxes is* EXPTIME*-complete.*

Further, if we drop role inclusions and consider global functionality, we show that, similar to Section 5, we can reduce satisfiability in *DL-Lite*$^{\mathcal{SF}}_{bool}$ to satisfiability in *DL-Lite*$^{\mathcal{F},\mathsf{sf}}_{bool}$, extending *DL-Lite*$^{\mathcal{F}}_{bool}$ with local reflexivity concepts. To obtain the desired result, we first show:

Lemma 4. *Concept satisfiability relative to DL-Lite*$^{\mathcal{F},\mathsf{sf}}_{bool}$ *TBoxes is* NP*-complete.*

The lower bound is inherited from *DL-Lite*$^{\mathcal{F}}_{bool}$. The upper bound can be proved by extending the reduction from *DL-Lite*$^{\mathcal{F}}_{bool}$ to the one-variable fragment of first-order logic (Artale et al. 2009) so as to deal with local reflexivity. With Lemma 4 at hand, we obtain the following (where the lower bound is also inherited from *DL-Lite*$^{\mathcal{F}}_{bool}$) :

Theorem 7. *Concept satisfiability relative to DL-Lite*$^{\mathcal{SF}}_{bool}$ *TBoxes is* NP*-complete.*

Note that the reduction from $DL\text{-}Lite_{bool}^{\mathcal{SF}}$ to $DL\text{-}Lite_{bool}^{\mathcal{F},\mathsf{sf}}$ relies on the availability of disjunction. Hence we cannot lift the reduction to non-Boolean complete fragments.

7 Conclusions and Future Work

In this paper, we have made progress on the understanding of the computational complexity of DLs allowing for number restrictions on transitive roles. In particular, we have established a tight NEXPTIME upper bound for satisifiability in $\mathcal{SOQ}$, and showed that in $\mathcal{SON}$ the coding of numbers plays a role on the computational complexity.

As the next step, we will look for ways to incorporate inverse roles and numbers greater than 1 without losing decidability. In this direction, we will investigate languages based on *DL-Lite without* RIs. In $\mathcal{SOIQ}$, decidability might be regained by admitting counting only over r or r^{-}, but not over both (Kazakov, Sattler, and Zolin 2007). We will also investigate ways to include some other forms of complex roles, such as role composition.

On the practical side, we are interested in developing a consequence-based calculus for our logics – a promising starting point is the recently proposed calculus for $\mathcal{SRIQ}$, supporting number restrictions on non-transitive roles (Bate et al. 2016).

We will also study DLs supporting counting over transitive roles in the context of *ontology-based data access*. We want to understand the impact of these features on the problem of *conjunctive query answering*, in the case where transitive roles occur in the query. Moreover, we will consider conjunctive queries incorporating some type of counting, e.g., restricted versions of inequalities, such as local inequalities (Gutiérrez-Basulto et al. 2015).

Acknowledgments We thank anonymous reviewers for useful comments and suggestions. The first author was supported by a Marie Curie SIRCIW COFUND grant.

References

Artale, A.; Calvanese, D.; Kontchakov, R.; and Zakharyaschev, M. 2009. The dl-lite family and relations. *J. Artif. Intell. Res. (JAIR)* 36:1–69.

Baader, F.; Lutz, C.; Sturm, H.; and Wolter, F. 2002. Fusions of description logics and abstract description systems. *J. Artif. Intell. Res. (JAIR)* 16:1–58.

Baader, F.; Calvanese, D.; McGuinness, D. L.; Nardi, D.; and Patel-Schneider, P. F., eds. 2003. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press.

Bate, A.; Motik, B.; Grau, B. C.; Simancik, F.; and Horrocks, I. 2016. Extending consequence-based reasoning to SRIQ. In *Proc. of KR-16*, 187–196.

Calvanese, D.; Eiter, T.; and Ortiz, M. 2009. Regular path queries in expressive description logics with nominals. In *Proc. of IJCAI-09*, 714–720.

Chagrov, A. V., and Zakharyaschev, M. 1997. *Modal Logic*, volume 35 of *Oxford logic guides*. Oxford University Press.

Gutiérrez-Basulto, V.; Ibáñez-García, Y. A.; Kontchakov, R.; and Kostylev, E. V. 2015. Queries with negation and inequalities over lightweight ontologies. *J. Web Sem.* 35:184–202.

Hollunder, B., and Baader, F. 1991. Qualifying number restrictions in concept languages. In *Proc. of KR-91*, 335–346.

Horrocks, I.; Sattler, U.; and Tobies, S. 2000. Practical reasoning for very expressive description logics. *Logic Journal of the IGPL* 8(3):239–263.

Kaminski, M., and Smolka, G. 2010. Terminating tableaux for $\mathcal{SOQ}$ with number restrictions on transitive roles. In *Theoretical Computer Science - 6th IFIP TC Proceedings*, 213–228.

Kazakov, Y., and Pratt-Hartmann, I. 2009. A note on the complexity of the satisfiability problem for graded modal logics. In *Proc. of LICS-09*, 407–416.

Kazakov, Y.; Sattler, U.; and Zolin, E. 2007. How many legs do I have? Non-simple roles in number restrictions revisited. In *Proc. of LPAR-07*, 303–317.

Lutz, C.; Areces, C.; Horrocks, I.; and Sattler, U. 2005. Keys, nominals, and concrete domains. *J. Artif. Intell. Res. (JAIR)* 23:667–726.

Lutz, C. 2004. An improved NExpTime-hardness result for the description logic $\mathcal{ALC}$ extended with inverse roles, nominals and counting. Technical Report LTCS-04-07, Chair for Automata Theory, TU Dresden, Germany.

Motik, B.; Shearer, R.; and Horrocks, I. 2009. Hypertableau reasoning for description logics. *J. Artif. Intell. Res. (JAIR)* 36:165–228.

Rector, A. L., and Rogers, J. 2006. Ontological and practical issues in using a description logic to represent medical concept systems: Experience from GALEN. In *Proc. of RW-06*, 197–231.

Schröder, L., and Pattinson, D. 2008. How many toes do I have? parthood and number restrictions in description logics. In *Proc. of KR-08*, 307–317.

Stevens, R.; Aranguren, M. E.; Wolstencroft, K.; Sattler, U.; Drummond, N.; Horridge, M.; and Rector, A. L. 2007. Using OWL to model biological knowledge. *International Journal of Man-Machine Studies* 65(7):583–594.

Tendera, L. 2005. Counting in the two variable guarded logic with transitivity. In *Proc. of STACS-05*, 83–96.

Tobies, S. 2001. *Complexity Results and Practical Algorithms for Logics in Knowledge Representation*. Ph.D. Dissertation, RWTH Aachen.

van Emde Boas, P. 1997. The convenience of tilings. In *Complexity, Logic, and Recursion Theory*, 331–363.

Wolstencroft, K.; Brass, A.; Horrocks, I.; Lord, P. W.; Sattler, U.; Turi, D.; and Stevens, R. 2005. A little semantic web goes a long way in biology. In *Proc. of ISWC-05*, 786–800.

APPENDIX

Proofs for Section 3

Proposition 1. *Every $\mathcal{SOQ}$-TBox $\mathcal{T}$ can be transformed in polynomial time into a TBox $\mathcal{T}'$ such that for all concept names A appearing in $\mathcal{T}$, we have that A is satisfiable relative to $\mathcal{T}$ iff A is satisfiable relative to $\mathcal{T}'$.*

Proof. Obtain a TBox $\mathcal{T}'$ from $\mathcal{T}$ as follows. Replace each $(\sim n\ r\ D)$ appearing in $\mathcal{T}$ by a fresh concept name $X_{(\sim n\ r\ D)}$ and include the concept inclusions

$$X_{(\sim n\ r\ D)} \sqsubseteq (\sim n\ r\ D)$$
$$\neg X_{(\sim n\ r\ D)} \sqsubseteq \begin{cases} (\leq n-1\ r\ D) & \text{if} \sim \text{is} \geq; \\ (\geq n+1\ r\ D) & \text{if} \sim \text{is} \leq. \end{cases}$$

It is routine to show that, for all concept names A appearing in $\mathcal{T}$, we have that A is satisfiable relative to $\mathcal{T}$ iff A is satisfiable relative to $\mathcal{T}'$. ☐

Theorem 1. *A is satisfiable relative to $\mathcal{T}$ iff there is a quasimodel for A and $\mathcal{T}$.*

Proof. For ($\Rightarrow$), take a model $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ of A and $\mathcal{T}$. For each role name $r \in \mathsf{Rol}(\mathcal{T})$, let $\mathcal{I}_r$ be obtained from $\mathcal{I}$ by setting $s^{\mathcal{I}_r} = \emptyset$ for all $s \neq r$. It is not hard to show that $\mathfrak{Q} = \{\mathcal{I}_r \mid r \in \mathsf{Rol}(\mathcal{T})\}$ is a quasimodel for A and $\mathcal{T}$.

For showing ($\Leftarrow$), take a quasimodel $\mathfrak{Q} = \{\mathcal{I}_r \mid r \in \mathsf{Rol}\}$ for A and $\mathcal{T}$. Assume without loss of generality that the domains of the interpretations coincide on the nominals and are disjoint otherwise. More formally, we assume $o^{\mathcal{I}_r} = o^{\mathcal{I}_s}$ for all $o \in \mathsf{Nom}(\mathcal{T})$ and $\Delta^{\mathcal{I}_r} \cap \Delta^{\mathcal{I}_s} = \{o^{\mathcal{I}_r} \mid o \in \mathsf{Nom}(\mathcal{T})\}$ for all role names $r \neq s$. We construct an interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ with the following domain and interpretation of concept names and nominals:

$$\Delta^{\mathcal{I}} = \bigcup_{r \in \mathsf{Rol}(\mathcal{T})} \Delta^{\mathcal{I}_r},$$
$$A^{\mathcal{I}} = \bigcup_{r \in \mathsf{Rol}(\mathcal{T})} A^{\mathcal{I}_r}, \quad \text{for all } A \in \mathsf{CN},$$
$$o^{\mathcal{I}} = o^{\mathcal{I}_r}, \quad \text{for some } \mathcal{I}_r.$$

Note that the interpretation of nominals is well-defined, due to the assumption on the domains. Moreover, it should be clear that we have for all $\mathcal{I}_r \in \mathfrak{Q}$, $d \in \Delta^{\mathcal{I}_r}$, and $C \in \mathcal{C}_{\mathsf{Bool}}$, we have that

$$d \in C^{\mathcal{I}_r} \quad \text{iff} \quad d \in C^{\mathcal{I}}. \tag{1}$$

For defining the interpretation of a role name r, we leverage Condition (**qm1**) to define a relation $\rightsquigarrow_r$ as follows. For each $d \in \Delta^{\mathcal{I}_s}$ with $s \neq r$, let $d' \in \Delta^{\mathcal{I}_r}$ the Boolean equivalent element in $\mathcal{I}_r$, which exists due to Condition (**qm2**). Then, include $d \rightsquigarrow_r e$ for all $(d', e) \in r^{\mathcal{I}_r}$. Intuitively, $d \rightsquigarrow_r e$ if the Boolean equivalent "copy" d' of d has the r-successor e. Then, define

$$r^{\mathcal{I}} = \rightsquigarrow_r \cup\, r^{\mathcal{I}^r}. \tag{†}$$

This finishes the definition of $\mathcal{I}$ and it remains to show that $\mathcal{I}$ is a model of A and $\mathcal{T}$. For the former, observe that $A^{\mathcal{I}} \neq \emptyset$, as by assumption, $A^{\mathcal{I}_r} \neq \emptyset$ for some r. For the latter, observe first that $\mathcal{I}$ satisfies concept inclusions of the form $C \sqsubseteq D$ for $C, D \in \mathcal{C}_{\mathsf{Bool}}$, due to (1). Consider now a concept inclusion $C \sqsubseteq (\sim n\ r\ D) \in \mathcal{T}$ and $d \in C^{\mathcal{I}}$. Moreover, assume that $d \in \Delta^{\mathcal{I}_s}$ for some role name s, thus $d \in C^{\mathcal{I}_s}$, by (1). We distinguish two cases.

- If $r = s$, then $d \in (\sim n\ r\ D)^{\mathcal{I}_s}$, since $C \sqsubseteq (\sim n\ r\ D) \in \mathcal{T}_s$ and $\mathcal{I}_s \models \mathcal{T}_s$, by Condition (**qm1**). Note that (†) does not add r-successors to d in this case, thus $d \in (\sim n\ r\ D)^{\mathcal{I}}$.
- If $r \neq s$, let $\mathcal{I}_r$ and $d' \in \Delta^{\mathcal{I}_r}$ as in Condition (**qm2**), that is, d, d' are Boolean equivalent. Thus, we also have that $d' \in C^{\mathcal{I}_r}$. As $C \sqsubseteq (\sim n\ r\ D) \in \mathcal{T}_r$ and $\mathcal{I}_r \models \mathcal{T}_r$, by Condition (**qm1**), we get $d \in (\sim n\ r\ D)^{\mathcal{I}_r}$. By definition of $\rightsquigarrow_r$ and $r^{\mathcal{I}}$, we obtain $d \in (\sim n\ r\ D)^{\mathcal{I}}$.

This finishes the proof of the lemma. ☐

Lemma 2. *If there is a quasimodel for A and $\mathcal{T}$, there is a quasimodel of exponential size.*

Let $\mathcal{I}_r \in \mathfrak{Q}$. In the main part, we have already captured the case when r is non-transitive. Moreover, we argued that, when r is transitive, every $\mathcal{I}_r$ in a quasimodel $\mathfrak{Q}$ consists of at most 2^X connected components. We show here that each such connected component can be assumed to be of exponential size. More precisely, let here and in what follows $(\geq n_i\ r\ D_i)$, $1 \leq i \leq \ell$ be all at-least restrictions appearing in $\mathsf{cl}(\mathcal{T})$, and set $b = \max(2, \sum_{i=1}^{\ell} n_i)$ and $d = 2\ell$. We show the following Lemma.

Lemma 5. *Let $\mathcal{I}$ be a connected interpretation such that $\mathcal{I} \models \mathcal{T}_r$ for some role name r. Then, there is an interpretation $\mathcal{I}'$ of size at most*

$$b^{d+2} + 2^{|X|}, \tag{$*$}$$

such that $\mathcal{I}' \models \mathcal{T}_r$ and for each $d \in \Delta^{\mathcal{I}}$ there is a $d' \in \Delta^{\mathcal{I}'}$ such that d, d' are Boolean equivalent, and vice versa.

This suffices since the number in ($*$) is bounded by some exponential, even though numbers are encoded in binary, and the latter property ensures that, we can replace $\mathcal{I}$ by $\mathcal{I}'$ in $\mathcal{I}_r$, c.f. condition (**qm2**).

To prove Lemma 5, we need some auxiliary notation. Denote with $r^*_{\mathcal{I}}(d, C)$ the extension of $r_{\mathcal{I}}(d, C)$ to the reflexive closure of $r^{\mathcal{I}}$. Moreover, denote with R^- and R^+ the inverse and transitive closure, respectively, of a binary relation R. We use the following terms:

- d' is an *r-successor* of d if $(d, d') \in r^{\mathcal{I}}$;
- d, d' are *r-equivalent* if $(d, d') \in r^{\mathcal{I}}$ and $(d', d) \in r^{\mathcal{I}}$;
- d' is *strict r-successor* of d if $(d, d') \in r^{\mathcal{I}}$ and $(d', d) \notin r^{\mathcal{I}}$;
- d' is a *direct r-successor* of d if d' is a strict r-successor of d and, for all e with $(d, e) \in r^{\mathcal{I}}$ and $(e, d') \in r^{\mathcal{I}}$, we have $(e, d) \in r^{\mathcal{I}}$ or $(d', e) \in r^{\mathcal{I}}$;
- the *r-cluster of* d, denoted $Q(d)$, is the set of all e such that both $(d, e) \in r^{\mathcal{I}}$ and $(e, d) \in r^{\mathcal{I}}$;
- the *depth of* $\mathcal{I}$ is the maximal number k such that there is a sequence $d_0, \ldots, d_k \in \Delta^{\mathcal{I}}$ such that d_{i+1} is a strict r-successor of d_i for all $0 \leq i < k$;

- the *breadth of* $\mathcal{I}$ is the maximal number k such that there are $d, d_0, \ldots, d_k$ such that $Q(d_i) \neq Q(d_j)$ for $i \neq j$ and d_i are direct r-successors of d;
- the *width of* $\mathcal{I}$ is the minimal number k such that $|Q(d)| \leq k$ for all $d \in \Delta^{\mathcal{I}}$ (∞ if no such k exists).

Proof. Let $\mathcal{I}$ be a connected component of some $\mathcal{I}_r \in \mathfrak{Q}$. We follow the 4-stage strategy from (Kazakov and Pratt-Hartmann 2009); the main difference is in stage 4, where we have to take care that a representative for every Boolean equivalence class is kept, c.f. the last condition in Lemma 5. Stage 1 also varies to correct an inaccurracy in (Kazakov and Pratt-Hartmann 2009).

Stage 1 (Finite depth). Let $(\leq m_j\ r\ C_j)$, $1 \leq j \leq k$ be all at-most restrictions occurring in $\mathcal{T}$. For each $d \in \Delta^{\mathcal{I}}$, we define a relation for each C_j, $1 \leq j \leq m$ as follows: $d \approx_j^{\mathcal{I}} d'$ if $(d, d') \in r^{\mathcal{I}}$ and either $|r_{\mathcal{I}}(d', C_j)| > m_j$ or $|r^*_{\mathcal{I}}(d, C_j)| = |r^*_{\mathcal{I}}(d', C_j)| \leq m_j$. Now, let

$$R := \{(d, d') \mid d \approx_j^{\mathcal{I}} d',\ \text{for every } 1 \leq j \leq k\}.$$

We obtain $\mathcal{I}'$ from $\mathcal{I}$ by setting $r^{\mathcal{I}'} = (r^{\mathcal{I}} \cup R^-)^+$, $A^{\mathcal{I}'} = A^{\mathcal{I}}$ for each $A \in \mathsf{CN}$, and $\{o\}^{\mathcal{I}} = \{o\}^{\mathcal{I}'}$, for every nominal $\{o\}$.

Claim 1. $C^{\mathcal{I}} = C^{\mathcal{I}'}$ for all $C \in \mathsf{cl}(\mathcal{T})$.

Proof of Claim 1. We show the claim by induction on the structure of concepts. Observe that both the base cases, $C = A$ a concept name or $C = \{o\}$ a nominal, and the Boolean cases, $C = \neg C_1$ or $C = C_1 \sqcap C_2$, are immediate. For $C = (\geq n_i\ r\ D_i)$, the claim follows from $r^{\mathcal{I}} \subseteq r^{\mathcal{I}'}$.

Consider next $C = (\leq m_j\ r\ C_j)$. We clearly have $C^{\mathcal{I}'} \subseteq C^{\mathcal{I}}$ since $r^{\mathcal{I}} \subseteq r^{\mathcal{I}'}$.

For the converse, we will show that for all d with $|r_{\mathcal{I}}(d, C_j)| \leq m_j$ it holds that

$$|r_{\mathcal{I}}(d, C_j)| = |r_{\mathcal{I}'}(d, C_j)| \tag{2}$$

This yields the required result. Indeed, assume $d \in C^{\mathcal{I}}$, thus $|r_{\mathcal{I}}(d, C_j)| \leq m_j$. Then, by (2), $|r_{\mathcal{I}'}(d, C_j)| \leq m_j$, and therefore $d \in (\leq m_j\ r\ C_j)^{\mathcal{I}'}$.

Then, suppose towards a contradiction that $|r_{\mathcal{I}'}(d, C_j)| > m_j$. This means that there is some $e \in C_j^{\mathcal{I}'}$ such that $(d, e) \in r^{\mathcal{I}'}$ and $(d, e) \notin r^{\mathcal{I}}$ or $e \notin C_j^{\mathcal{I}}$. By induction hypothesis, whe know that $e \in C_j^{\mathcal{I}}$, then it must be the case that $(d, e) \notin r^{\mathcal{I}}$. By the definition of $r^{\mathcal{I}'}$, this means that there exists a sequence $d = d_0, \ldots, d_k = e \in \Delta^{\mathcal{I}}$ such that $d_0 = d$, $(d_i, d_{i+1}) \in r^{\mathcal{I}} \cup R^-$ for all $0 \leq i < k$. Take the maximal i such that $(d_i, e) \notin r^{\mathcal{I}}$. As $(d, e) \notin r^{\mathcal{I}}$, such maximal i always exists. Then $(d_{i+1}, e) \in (r^{\mathcal{I}})^*$ and $(d_i, d_{i+1}) \notin r^{\mathcal{I}}$. Since $(d_i, d_{i+1}) \in r^{\mathcal{I}} \cup R^-$, we have $(d_i, d_{i+1}) \in R^-$ and, by definition of R, $(d_{i+1}, d_i) \in r^{\mathcal{I}}$ and $d_i \approx_j^{\mathcal{I}} d_{i+1}$. Note that $(d, d_i) \in r^{\mathcal{I}}$, and $d \in (\leq m_j\ r\ C_j)$ imply that $|r_{\mathcal{I}}(d_i, C_j)| \leq m_j$. Then, $|r^*_{\mathcal{I}}(d_i, C_j)| = |r^*_{\mathcal{I}}(d_{i+1}, C_j)| \leq m_j$, as $d_i \approx_j^{\mathcal{I}} d_{i+1}$. This in turn implies that $e \notin C_j^{\mathcal{I}}$, a contradiction, as otherwise we would have $|r^*_{\mathcal{I}}(d_{i+1}, C_j)| \geq |r^*_{\mathcal{I}}(d_i, C_j)| + 1$, since $(d_i, e) \notin r^{\mathcal{I}}$.

It remains to show that $\mathcal{I}'$ has finite depth. For $d \in \Delta^{\mathcal{I}}$ define $w_j^{\mathcal{I}}(d) := \min(m_j + 1, |r^*_{\mathcal{I}}(d, C_j)|)$. Note that for $d, d' \in \Delta^{\mathcal{I}}$, $d \approx_j^{\mathcal{I}} d'$ implies $w_j^{\mathcal{I}}(d) = w_j^{\mathcal{I}}(d')$[4]. Then for every $d_1, d_2 \in \Delta^{\mathcal{I}'}$ such that d_2 is a strict successor of d_1 (in $\mathcal{I}'$), we have $w_j^{\mathcal{I}}(d_1) \geq w_j^{\mathcal{I}}(d_2)$ for all j; and $w_j^{\mathcal{I}}(d_1) > w_j^{\mathcal{I}}(d_2)$ for some j ($1 \leq j \leq k$). Hence, $\sum_{j=1}^{k} w_j^{\mathcal{I}}(d_1) > \sum_{j=1}^{k} w_j^{\mathcal{I}}(d_2)$. Since $w_j^{\mathcal{I}}(d) \leq m_j + 1$ for every $d \in \Delta^{\mathcal{I}}$ and every j, the lenght of every chain $d_0, \ldots, d_l$ with d_{i+1} a strict successor of d_i ($0 \leq i < l$), is bounded by $\sum_{j=1}^{k} m_j + k$.

Stage 2 (Bounded Breadth). By stage 1, we can assume that $\mathcal{I}$ has finite depth. We define an r-witness $\mathcal{I}'$ of finite depth and breadth $b' \leq \sum_{i=1}^{\ell} n_i$.

For every element $d \in \Delta^{\mathcal{I}}$ and every $1 \leq i \leq \ell$, let $W_i(d)$ be the set of strict r-successors satisfying D_i. Note that $W_i(d_1) = W_i(d_2)$ if d_1 and d_2 are r-equivalent. Let $W_i'(d)$ be $W_i(d)$ if $|W_i(d)| \leq n_i$, and a subset of $W_i(d)$ of size exactly n_i, otherwise. We assume wlog that $W_i'(d_1) = W_i'(d_2)$ if d_1 and d_2 are r-equivalent. Define relations R_q and R_i' as follows:

$$R_q = \{(d, d') \in r^{\mathcal{I}} \mid d' \in Q(d)\};$$
$$R_i' = \{(d, d') \in r^{\mathcal{I}} \mid w' \in W_i'(d)\}.$$

Intuitively, R_q is the restriction of $r^{\mathcal{I}}$ to the cliques and R_i' keeps only those successors required to witness the concept D_i from the i-th at-least restriction.

Finally, obtain $\mathcal{I}'$ by taking $\Delta^{\mathcal{I}'} = \Delta^{\mathcal{I}}$, $A^{\mathcal{I}'} = A^{\mathcal{I}}$ for all concept names A, and

$$r^{\mathcal{I}'} = (R_q \cup \bigcup_{1 \leq i \leq \ell} R_i')^+.$$

Clearly, $r^{\mathcal{I}'}$ is transitive and has breadth b' as required.

Claim 3. $C^{\mathcal{I}} = C^{\mathcal{I}'}$ for all $C \in \mathsf{cl}(\mathcal{T})$.

Proof of Claim 3. This is again shown by induction on the structure of concepts. The only non-trivial case are concepts $C = (\geq n_i\ r\ D_i)$. Clearly, $d \in C^{\mathcal{I}'}$ implies $d \in C^{\mathcal{I}}$ since $r^{\mathcal{I}'} \subseteq r^{\mathcal{I}}$. The converse is a direct consequence of the definition of $W_i'(d)$ and R_i'. In particular, only r-connections that are not necessary for witnessing the at least-restrictions are removed.

Stage 3 (Bounded Depth). By stage 2, we can assume that $\mathcal{I}$ has finite depth d and finite breadth $b \leq \sum_{i=1}^{\ell} n_i$. We define an r-witness $\mathcal{I}$' of depth $d' \leq 2\ell$. If already $d \leq 2\ell$, we are done. For the other case, we define an operation $\overline{\cdot}$ that, if applied exhaustively, establishes the mentioned goal. For every $d \in \Delta^{\mathcal{I}}$, define two sets of concepts:

$$X(d) = \{D_i \mid 1 \leq i \leq \ell \text{ and } |r_{\mathcal{I}}(d, D_i)| \geq n_i\}, \text{ and}$$
$$X'(d) = \{D_i \mid 1 \leq i \leq \ell \text{ and } |r_{\mathcal{I}}(d, D_i) \setminus Q(d)| \geq n_i\}.$$

[4]The converse, however, does not hold in general. This seems to be the overlook in (Kazakov and Pratt-Hartmann 2009)

Fix now two elements d, d' such that d' is a direct r-successor of d and $X'(d') = X(d)$, and obtain $\overline{\mathcal{I}}$ by taking $\Delta^{\overline{\mathcal{I}}} = \Delta^{\mathcal{I}}$, $A^{\overline{\mathcal{I}}} = A^{\mathcal{I}}$ for all A, and

$$r^{\overline{\mathcal{I}}} = r^{\mathcal{I}} \setminus (Q(d) \times Q(d'))$$

We first show the following two properties of $\overline{\mathcal{I}}$.

Claim 4. $C^{\mathcal{I}} = C^{\overline{\mathcal{I}}}$ for all $C \in \mathsf{cl}(\mathcal{T})$.

Proof of Claim 4. This is again shown by induction on the structure of concepts. The only non-trivial case are concepts $C = (\geq n_i\ r\ D_i)$. Clearly, $d \in C^{\overline{\mathcal{I}}}$ implies $d \in C^{\mathcal{I}}$ since $r^{\overline{\mathcal{I}}} \subseteq r^{\mathcal{I}}$. The converse is a direct consequence of the definition of $r^{\overline{\mathcal{I}}}$, in particular, we remove only r-connections that are not necessary for witnessing the at least-restrictions.

Claim 5. $r^{\overline{\mathcal{I}}}$ is transitive.

Proof of Claim 5. Assume d, e, f with $(d, e) \in r^{\overline{\mathcal{I}}}$ and $(e, f) \in r^{\overline{\mathcal{I}}}$, but $(d, f) \notin r^{\overline{\mathcal{I}}}$. The former implies $(d, e) \in r^{\mathcal{I}}$ and $(e, f) \in r^{\mathcal{I}}$, thus $(d, f) \in r^{\mathcal{I}}$, by transitivity of $r^{\mathcal{I}}$. As $(d, f) \notin r^{\overline{\mathcal{I}}}$, we know that $X'(f) = X(d)$ and f a direct r-successor of d, thus $e \in Q(d)$ or $e \in Q(f)$. In the former case, f is a direct successor of e, and $X(d) = X(e)$. Thus $X(e) = X'(f)$, and $(e, f) \notin r^{\overline{\mathcal{I}}}$. In the latter case, we argue analogously. This finishes the proof of Claim 5.

Now, let $\mathcal{I}'$ be obtained from $\mathcal{I}$ by exhaustively applying $\overline{\cdot}$. Since $\mathcal{I}$ is of finite depth and breadth (by Stages 1 and 2), this is a finite process; thus, $\mathcal{I}'$ satisfies Claims 4 and 5 and can replace $\mathcal{I}$. It remains to verify:

Claim 6. The depth d' of $\mathcal{I}'$ is $d' \leq 2\ell$.

Proof of Claim 6. Suppose the contrary, that is, there are $k > 2\ell$ elements $d_0, \ldots, d_k$ elements in $\Delta^{\mathcal{I}'}$ such that d_{i+1} is a direct r-successor of d_i for all $0 \leq i < k$. As d_{i+1} is a direct r-successor of d_i, we know

$$r_{\mathcal{I}}(d_{i+1}, D) \subseteq r_{\mathcal{I}}(d_i, D) \setminus Q(d_i), \quad \text{for all } i \text{ and } D, \quad (3)$$

which implies that $X(d_{i+1}) \subseteq X'(d_i)$ for all $0 \leq i < k$. By definition of X and X', we have that $X'(d_i) \subseteq X(d_i)$, for all

$$X(d_k) \subseteq X'(d_{k-1}) \subseteq X(d_{k-1}) \subseteq \ldots$$
$$\ldots \subseteq X(d_1) \subseteq X'(d_0) \subseteq X(d_0).$$

As $k > 2\ell$, we find i such that $X(d_i) = X'(d_{i+1})$. Thus, by construction of $r^{\mathcal{I}'}$, the edge between d_i and d_{i+1} was removed, contradiction. This finishes the proof of Claim 6, and in fact of Stage 3.

Stage 4 (Bounded Model). By stage 3, we can assume that $\mathcal{I}$ has depth $d \leq 2\ell$ and breadth $b \leq \sum_{i=1}^{\ell} n_i$. We finally restrict the size of the domain of $\mathcal{I}$, by showing that we can replace $\mathcal{I}$ by an interpretation $\mathcal{I}' = (\Delta^{\mathcal{I}'}, \cdot^{\mathcal{I}'})$ of size at most $(*)$ as required.

For doing so, fix for every $Y \subseteq X$, domain elements d_Y such that $Y = \{C \in X \mid d_Y \in C^{\mathcal{I}_r}\}$ (if such an element exists), and let Δ_0 be the set of all the fixed elements.

Then, let $Q_i(d)$ be the subset of $Q(d)$ satisfying D_i, for all $1 \leq i \leq \ell$. Note that $Q_i(d_1) = Q_i(d_2)$ whenever d_1, d_2 are r-equivalent. Let $Q_i'(d)$ be $Q_i(d)$ if $|Q_i(d)| \leq n_i$, and a subset of $Q_i(d)$ with exactly n_i elements otherwise. Assume w.l.o.g. that $Q_i'(d_1) = Q_i'(d_2)$ whenever d_1 and d_2 are r-equivalent.

Now, define the domain of $\mathcal{I}'$ as

$$\Delta^{\mathcal{I}'} = \bigcup_{d \in \Delta^{\mathcal{I}}, 1 \leq i \leq \ell} Q_i'(d) \cup \Delta_0,$$

and $A^{\mathcal{I}'} = A^{\mathcal{I}} \cap \Delta^{\mathcal{I}'}$ for all concept names A, and $r^{\mathcal{I}'} = r^{\mathcal{I}} \cap (\Delta^{\mathcal{I}'} \times \Delta^{\mathcal{I}'})$. It should be clear that $r^{\mathcal{I}'}$ is indeed transitive. To see that $\Delta^{\mathcal{I}'}$ has size at most $(*)$ note that, without Δ_0, it can be viewed as a b-ary tree of depth d, where each node corresponds to a clique of size at most b domain elements. Such a tree has $\leq b^{d+1}$ nodes, so there are at most b^{d+2} domain elements. The remaining constant $2^{|X|}$ then comes from the inclusion of Δ_0.

Finally note that the inclusion of Δ_0 implies that witnesses for the Boolean equivalence are kept. It remains to show that $\mathcal{I}' \models \mathcal{T}_r$, which is a consequence of the following

Claim 7. for all $C \in \mathsf{cl}(\mathcal{T})$ and all $d \in \Delta^{\mathcal{I}'}$, $d \in C^{\mathcal{I}}$ iff $d \in C^{\mathcal{I}'}$.

Proof of Claim 7. This is again shown by induction on the structure of concepts. The only non-trivial case are concepts $C = (\geq n_i\ r\ D_i)$. Clearly, $d \in C^{\mathcal{I}'}$ implies $d \in C^{\mathcal{I}}$ since $r^{\mathcal{I}'} \subseteq r^{\mathcal{I}}$. The converse is a direct consequence of the definition of $Q_i'(d)$ and $\Delta^{\mathcal{I}'}$. In particular, we drop only elements from clusters such that all at-least restrictions that were satisfied by this cluster are still satisfied. □

Proofs for Section 4.1

We directly prove Lemma 3 with (**E3**) replaced by (**E3'**).

Lemma 3. *A is satisfiable relative to $\mathcal{T}$ iff there is set $T \subseteq \mathsf{tp}(\mathcal{T})$ of types with $A \in t$ for some $t \in T$ such that:*

(E1) *for any $\{o\}$ in $\mathcal{T}$ there is exactly one $t \in T$ with $\{o\} \in t$;*

(E2) *every $t \in T$ is r-realisable in T, for each role r;*

(E3') *there is an interpretation $\mathcal{I}_r$ such that, for each $t \in T$ with $(\leq n\ r) \in t$, there is a $\leq$-witness $\mathcal{I}_{rt}$ for $(r, T, \{t\})$ with $\mathsf{core}_r(\mathcal{I}_{rt}) = \mathcal{I}_r$.*

Proof. The "only if"-direction is straightforward. Given a model $\mathcal{J}$ of A and $\mathcal{T}$, we can read off T as the set types realized by $\mathcal{J}$. It is routine to show that T satisfies (**E1**) and (**E2**). For (**E3'**), fix some transitive role r, and a type $t \in T$ with $(\leq n\ r) \in t$. We read off an $\leq$-witnesss $\mathcal{I}_{rt}$ for $(r, T, \{t\})$ by fixing an element d_t with $\mathsf{tp}_{\mathcal{I}}(d) = t$ and restricting $\mathcal{J}$ to the domain

$$\Delta^{\mathcal{I}_t} = \{o^{\mathcal{I}} \mid \{o\} \in \mathsf{cl}(\mathcal{T})\} \cup \{e \mid (d_t, e) \in (r^{\mathcal{I}})^*\}.$$

It is not hard to verify that these $\mathcal{I}_t$ together satisfy (**E3'**); in particular, all these $\mathcal{I}_t$ have an isomorphic nominal core $\mathcal{I} = \mathsf{core}_r(\mathcal{I}_t)$.

For the "if"-direction, assume some set T satisfying (**E1**)-(**E3'**), and assume that $\hat{A} \in \hat{t}$ for some $\hat{t} \in T$. By (**E1**), for each nominal $\{o\} \in \mathsf{cl}(\mathcal{T})$, there is a unique type t_o with $\{o\} \in t$. We construct a model $\mathcal{I}$ of A and $\mathcal{T}$. Define an interpretation $\mathcal{I}_0$ by taking:

$$\Delta^{\mathcal{I}_0} = \{d_t \mid t \in T\},$$
$$A^{\mathcal{I}_0} = \{d_t \mid A \in t\}, \qquad \text{for all concept names } A,$$
$$r^{\mathcal{I}_0} = \emptyset, \qquad \text{for all role names } r,$$
$$o^{\mathcal{I}_0} = d_{t_o}.$$

Additionally, initialize a map $\pi : \Delta^{\mathcal{I}_0} \to T$ by taking $\pi(d_t) = t$. Obtain an interpretation $\mathcal{I}_{i+1}$ from $\mathcal{I}_i$ as follows. Choose a domain element d with $\pi(d) = t$ such that $\exists r.C \in t$ but $d \notin (\exists r.C)^{\mathcal{I}_i}$, or $(\geq n\ r) \in t$, but $d \notin (\geq n\ r)^{\mathcal{I}_i}$. Denote with u the minimal number such that $(\leq u\ r) \in \pi(d)$, or ∞ if no such number exists. We distinguish cases.

If r is non-transitive or r is transitive and $u = \infty$, we proceed as follows. By (**E2**), t is r-realizable, that is, there are types $t_1, \ldots, t_k \in T$ which satisfy $t \rightsquigarrow_r t_i$, for all i and conditions 1-2. Let ℓ be as in the realizability condition, and suppose wlog that, for some N, $t_1, \ldots, t_N$ all contain some $\{o\} \in \mathsf{cl}(\mathcal{T})$ and $t_{N+1}, \ldots, t_k$ do not. Now, add fresh domain elements $d_{N+1}, \ldots, d_k$ and set $\pi(d_i) = t_i$, for each $N + 1 \leq i \leq k$. If $k < \ell$, then add $\ell - k$ fresh elements and set $\pi(d) = t_k$, for each of them (t_k does not contain a nominal, as $N < k$ by Item 2 of the realizability condition). Then, add (d, e) to $r^{\mathcal{I}_{i+1}}$, for all freshly introduced individuals e. Finally, take the transitive closure of $r^{\mathcal{I}_{i+1}}$, in case r is transitive.

If, on the other hand, r is transitive and $u < \infty$, then by (**E3'**), there is a $\leq$-witness $\mathcal{I}_{rt}$ for $(r, T, \{t\})$, that is, for each $d \in \Delta^{\mathcal{I}_{rt}}$, there is some type $\kappa(d) \in T$ with $\mathsf{tp}_{\mathcal{I}_{rt}}(d) =_r \kappa(d)$.

1. Suppose $t = t_o$ for some $\{o\} \in \mathsf{cl}(\mathcal{T})$. If there is a type $t_{o'}$, $\{o'\} \in \mathsf{cl}(\mathcal{T})$, such that $\{\exists r.\{o\}, (\leq n'\ r)\} \subseteq t_{o'}$, then do nothing. Otherwise, let D be the set of all $e \in \Delta^{\mathcal{I}_{rt}}$ such that $(o^{\mathcal{I}_{rt}}, e) \in r^{\mathcal{I}_{rt}}$ and $e \neq o'^{\mathcal{I}_{rt}}$, for every $\{o'\} \in \mathsf{cl}(\mathcal{T})$. Now, for each $e \in D$, add a fresh domain element e' to $\mathcal{I}_{i+1}$ and set $\pi(e') = \kappa(e)$. Finally, set $(e', f') \in r^{\mathcal{I}_{i+1}}$ iff $(e, f) \in r^{\mathcal{I}_{rt}}$, and $(\overline{o}^{\mathcal{I}_i}, e') \in r^{\mathcal{I}_{i+1}}$ iff $(\overline{o}^{\mathcal{I}_{rt}}, e) \in r^{\mathcal{I}_{rt}}$, for all $\{\overline{o}\} \in \mathsf{cl}(\mathcal{T})$.
2. Otherwise $\{o\} \notin t$ for all $\{o\} \in \mathsf{cl}(\mathcal{T})$. If there is some $\{o'\} \in \mathsf{cl}(\mathcal{T})$ with $\exists r.\{o\} \in t_{o'}$ and $(\leq n\ r) \in t_{o'}$, then do nothing. Otherwise, let D be the set of all $e \in \Delta^{\mathcal{I}_{rt}}$ such that $(d_0, e) \in r^{\mathcal{I}_{rt}}$ and there is no $\{o\} \in \mathsf{cl}(\mathcal{T})$ such that $(d_0, o^{\mathcal{I}_{rt}}), (o^{\mathcal{I}_{rt}}, e) \in r^{\mathcal{I}_{rt}}$. Now, for each $e \in D$, add a fresh domain element e' to $\mathcal{I}_{i+1}$ and set $\pi(e') = \kappa(e)$. Finally, set $(d, e') \in r^{\mathcal{I}_{i+1}}$ iff $(d_0, e) \in r^{\mathcal{I}_{rt}}$, $(e', f') \in r^{\mathcal{I}_{i+1}}$ iff $(e, f) \in r^{\mathcal{I}_{rt}}$, $(e', o^{\mathcal{I}_{i+1}}) \in r^{\mathcal{I}_{i+1}}$ iff $(e, o^{\mathcal{I}_{rt}}) \in r^{\mathcal{I}_{rt}}$.

In both cases, take the transitive closure of $r^{\mathcal{I}_{i+1}}$. It remains to give the interpretation of concept names in $\mathcal{I}_{i+1}$:

$$A^{\mathcal{I}_{i+1}} = A^{\mathcal{I}_i} \cup \{e \in \Delta^{\mathcal{I}_{i+1}} \setminus \Delta^{\mathcal{I}_i} \mid A \in \pi(e)\}.$$

Obtain the interpretation $\mathcal{I}$ as the limit, that is, $\Delta^{\mathcal{I}} = \bigcup_{i \geq 0} \Delta^{\mathcal{I}_i}$, $A^{\mathcal{I}} = \bigcup_{i \geq 0} A^{\mathcal{I}_i}$, for all $A \in \mathsf{CN}$, and $r^{\mathcal{I}} = \bigcup_{i \geq 0} r^{\mathcal{I}_i}$ for all role names r.

Claim 1. $(d, e) \in r^{\mathcal{I}}$ implies $\pi(d) \rightsquigarrow_r \pi(e)$.

Proof of Claim 1. The Claim is trivially true in $\mathcal{I}_0$. We distinguish cases to show that the property is preserved when going from $\mathcal{I}_i$ to $\mathcal{I}_{i+1}$:

- If r is non-transitive or r is transitive and $u = \infty$, then, by the realizability condition, only r-compatible elements are connected in the first step. It remains to note that the property is preserved under taking transitive closure, since $\rightsquigarrow_r$ itself is transitive for transitive roles r.
- if r is transitive and $u \neq \infty$, it suffices to note that the introduced labeling κ of an $\leq$-witness satisfies the property and that, again, the property is preserved under taking transitive closure.

This finishes the proof of Claim 1.

Claim 2. $d \in C^{\mathcal{I}}$ iff $C \in \pi(d)$ for all $C \in \mathsf{cl}(\mathcal{T})$.

Proof of the Claim 2. We prove the claim by induction on the structure of concepts. It is certainly true for all concept names $A \in \mathsf{cl}(\mathcal{T})$, by definition of $A^{\mathcal{I}}$. The cases of $C = \neg D$ and $C = D_1 \sqcap D_2$ follow directly from the induction hypothesis. It remains to treat the cases $C = \exists r.D$ and $C = (\sim n\ r)$. We start with the former.

For the "if"-direction, it should be clear that the construction ensures that, if $(\exists r.D) \in \pi(d)$, also $d \in (\exists r.D)^{\mathcal{I}}$.

For the "only if"-direction, suppose $d \in (\exists r.D)^{\mathcal{I}}$, that is, there is $e \in D^{\mathcal{I}}$ such that $(d, e) \in r^{\mathcal{I}}$. By Claim 1, we know that $\pi(d) \rightsquigarrow_r \pi(e)$. By induction we know that $D \in \pi(e)$. The definition of $\rightsquigarrow_r$ then yields $\exists r.D \in \pi(d)$.

For $(\sim n\ r)$, it suffices to prove the "if"-direction for both $\leq$ and $\geq$, since $(\leq n\ r) \notin t$ implies $(\geq n + 1\ r) \in t$ and $(\geq n\ r) \notin t$ implies $(\leq n - 1\ r) \in t$.

Suppose first $(\geq n\ r) \in \pi(d)$. It should be clear that the construction ensures that then $d \in (\geq n\ r)^{\mathcal{I}_i}$ for some i and thus $d \in (\geq n\ r)^{\mathcal{I}}$, as no connections are removed.

Suppose now $(\leq n\ r) \in \pi(d)$ and denote with u the minimal number such that $(\leq u\ r) \in \pi(d)$. By definition of $r^{\mathcal{I}_0}$, we have $d \in (\leq u\ r)^{\mathcal{I}_0}$. Observe next that the construction rule is applied at most once to each domain element d. If r is non-transitive, it has at most u r-successors after the application of this step. For the case when r is transitive, assume $t = \pi(d)$. It is crucial to observe that, by our construction and (**E3'**), $\mathcal{I}$ rooted at d is isomorphic to $\mathcal{I}_t$, rooted at that element $\hat{d}$ which satisfies $\mathsf{tp}_{\mathcal{I}_t}(\hat{d}) =_r t$. By the properties of the $\leq$-witness, we know that $\hat{d} \in (\leq n\ r)^{\mathcal{I}_{rt}}$, and thus $d \in (\leq n\ r)^{\mathcal{I}}$, by the isomorphism. This finishes the proof of the Claim.

It follows that $\mathcal{I} \models \mathcal{T}$: suppose $C \sqsubseteq D \in \mathcal{T}$ and $d \in C^{\mathcal{I}}$. By the Claim, $C \in \pi(d)$. Since $\pi(d)$ is a type and $C \sqsubseteq D \in \mathcal{T}$, we obtain $D \in \pi(d)$. Thus $d \in D^{\mathcal{I}}$, by the Claim again. Moreover, $\hat{A}^{\mathcal{I}} \neq \emptyset$ because $d_{\hat{t}} \in \hat{A}^{\mathcal{I}_0}$, and thus $d_{\hat{t}} \in \hat{A}^{\mathcal{I}}$. □

Proofs for Section 4.2

To complete the reduction started in the main part, we formally introduce tiling problems.

A *tiling system* is a triple $D = (T, H, V)$, where T is a finite set of *tile types* and $H, V \subseteq T \times T$ represent the *horizontal* and *vertical matching conditions*. An *initial condition for D* is of the form $\vec{t} = (t_0, \dots t_{n-1}) \in T^n$. A mapping $\tau : \{0, \dots, 2^n - 1\} \times \{0, \dots, 2^n - 1\} \to T$ is a *solution* for $(D, \vec{t})$ iff for all $x, y \leq 2^n$ the following holds:

– if $\tau(x, y) = t$ and $\tau(x \oplus_{2^n} 1, y) = t'$, then $(t, t') \in H$;

– if $\tau(x, y) = t$ and $\tau(x, y \oplus_{2^n} 1) = t'$, then $(t, t') \in V$;

– $\tau(i, 0) = t_i$, for $0 \leq i < n$.

We say that D *tiles the* $2^n \times 2^n$*-torus with initial condition* $\vec{t}$ in case there is a solution for $(D, \vec{t})$.

It is well-known that the following is a NExpTime-complete problem (van Emde Boas 1997): Given a tiling system D and an initial condition $\vec{t} = t_0, \dots, t_{n-1}$, does D tile the $2^n \times 2^n$-torus with initial condition $\vec{t}$?

To finish the reduction, fix a tiling system D and an initial condition $\vec{t}$ of length n. The TBox $\mathcal{T}_{\mathsf{tile}}$ contains $\mathcal{T}_{\mathsf{tor}}$, as constructed in the main part, and further concept inclusions using

- fresh concept names A_t, A'_t for each tile type in $t \in T$;
- for each $0 \leq i < n$, the abbreviation B_i which holds precisely at those d with $\mathsf{xpos}(d) = i$ (this is easily realized by a conjunction of the X_j, $\neg X_j$ according to the binary encoding of i).

We add the following concept inclusions to $\mathcal{T}_{\mathsf{tile}}$ to enforce that every element in level $2n$ satisfies a unique tile type, and that the initial condition is satisfied.

$$L_{2n} \sqsubseteq \bigsqcup_{t \in T} A_t \sqcap \prod_{t, t' \in T, t \neq t'} \neg(A_t \sqcap A_{t'}) \quad (4)$$

$$L_{2n} \sqcap \prod_{i=0}^{n-1} \neg Y_i \sqsubseteq \prod_{i=0}^{n-1} (B_i \to A_{t_i}) \quad (5)$$

We use the glueing points to ensure the compatibility conditions. In particular, we propagate A_t (or a copy A'_t) to the horizontal and vertical successors and check compatibility with H, V using the following concept inclusions:

$$\begin{aligned}
L_{2n} \sqcap A_t \sqcap X_i^* &\sqsubseteq \forall r.((H \sqcap X_i^*) \to A'_t) && \text{for } t \in T, \\
L_{2n} \sqcap A_t \sqcap X_i^* &\sqsubseteq \forall r.((H \sqcap X_i^+) \to A_t) && \text{for } t \in T, \\
L_{2n} \sqcap A_t \sqcap Y_i^* &\sqsubseteq \forall r.((V \sqcap Y_i^*) \to A'_t) && \text{for } t \in T, \\
L_{2n} \sqcap A_t \sqcap Y_i^* &\sqsubseteq \forall r.((V \sqcap Y_i^+) \to A_t) && \text{for } t \in T, \\
H \sqcap A_t \sqcap A'_{t'} &\sqsubseteq \bot && \text{for } (t, t') \notin H, \\
V \sqcap A_t \sqcap A'_{t'} &\sqsubseteq \bot && \text{for } (t, t') \notin V.
\end{aligned}$$

This finishes the construction of $\mathcal{T}_{\mathsf{tile}}$. It remains to show:

Lemma 6. *L_0 is satisfiable relative to $\mathcal{T}_{\mathsf{tile}}$ iff D tiles the $2^n \times 2^n$-torus with initial condition $\vec{t}$.*

Proof. The ($\Leftarrow$)-direction is straightforward.

For the ($\Rightarrow$)-direction, take any model $\mathcal{I}$ of L_0 and $\mathcal{T}_{\mathsf{tile}}$. Since $\mathcal{I} \models \mathcal{T}_{\mathsf{tor}}$, one can easily show, using the intuitions provided in the main part, that $\mathcal{I}$ "contains" a $2^n \times 2^n$-torus in the following sense:

1. the torus points P are precisely the elements from $L_{2n}^{\mathcal{I}}$, and each $d \in P$ has coordinates $\mathsf{pos}(d)$;
2. for each $0 \leq i, j < 2^n$, there is a point $d \in P$ with $\mathsf{pos}(d) = (i, j)$;
3. if $\mathsf{pos}(d) = \mathsf{pos}(d')$, then $d = d'$, for all $d, d' \in P$;
4. for any two points $d, d' \in P$ with $\mathsf{pos}(d) = (x, y)$ and $\mathsf{pos}(d') = (x', y')$ we have:
   – $x' = x \wedge y' = y \oplus_{2^n} 1$ iff there is $g \in V^{\mathcal{I}}$ with $(d, g), (d', g) \in r^{\mathcal{I}}$ and $\mathsf{pos}(g) = (x', y')$;
   – $x' = x \oplus_{2^n} 1 \wedge y' = y$ iff there is $g \in H^{\mathcal{I}}$ with $(d, g), (d', g) \in r^{\mathcal{I}}$ and $\mathsf{pos}(g) = (x', y')$;

By Items 1-3 above, we can associate a unique point $p_{ij} \in P$ with each pair $(i, j) \in \{0, \dots, 2^n - 1\}^2$ such that $\mathsf{pos}(p_{ij}) = (i, j)$. We read off a solution of the tiling system by defining $\tau(i, j)$ as the unique $t \in T$ such that $p_{ij} \in A_t^{\mathcal{I}}$. Uniqueness is implied by (4). By (5), τ satisfies the initial condition.

For the compatibility relation, suppose $d, d' \in P$ with $\mathsf{pos}(d) = (x, y)$ and $\mathsf{pos}(d') = (x \oplus_{2^n} 1, y')$, and let $\tau(x, y) = t$ and $\tau(x \oplus_{2^n} 1, y') = t'$, that is $d \in A_t^{\mathcal{I}}$ and $d' \in A_{t'}^{\mathcal{I}}$. By Item 4 above, there is some point $g \in H^{\mathcal{I}}$ with $(d, g), (d', g) \in r^{\mathcal{I}}$ and $\mathsf{pos}(g) = (x \oplus_{2^n} 1, y')$. By the last set of concept inclusions, we know that $g \in A_t^{\mathcal{I}}$ and $g \in {A'_{t'}}^{\mathcal{I}}$. Since $\mathcal{I} \models \mathcal{T}$, we know that $(t, t') \in H$.

The same arguments apply for the vertical compatibility, which finishes the proof. □

Proofs for Section 5

Lemma 7. *For every concept name A, A is satisfiable relative to $\mathcal{T}$ iff A is satisfiable relative to $\mathcal{T}'$.*

Proof. We start by introducing auxiliary notions used from here on. A role r is *simple wrt. a TBox $\mathcal{T}$*, if there is no role s with $s \sqsubseteq^* r$ such that $\mathsf{Tra}(s) \in \mathcal{T}$. A role is *non-simple wrt. $\mathcal{T}$* if is *not* simple wrt. $\mathcal{T}$. If the TBox is clear from the context we just say that a role r is (non-)simple.

For the ($\Rightarrow$)-direction, let $\mathcal{I}$ be a model of A and $\mathcal{T}$. We will show that $\mathcal{I}$ is also a model of $\mathcal{T}'$. First observe that for every role r such that $\mathsf{Tra}(r) \in \mathcal{T}$ and every $C \in \mathsf{cl}(\mathcal{T})$, we have $\mathcal{I} \models \forall r.C \sqsubseteq \forall r.(\forall r.C)$, by transitivity of r. It remains to show that for every such role r, $\mathcal{I}$ is also a model of $(\leq 1\ r) \sqsubseteq \forall r.(\forall r.\bot \sqcup \exists r.\mathsf{self})$. Indeed, let $d \in (\leq 1\ r)^{\mathcal{I}}$. Assume there is some $d' \in \Delta^{\mathcal{I}}$ such that $(d, d') \in r^{\mathcal{I}}$ (if no such d' exists, the claim holds trivially). We aim at showing that $d' \in (\forall r.\bot \sqcup \exists r.\mathsf{self})^{\mathcal{I}}$. If there is no $d'' \in \Delta^{\mathcal{I}}$ such that $(d', d'') \in r^{\mathcal{I}}$, then clearly, $d' \in (\forall r.\bot)^{\mathcal{I}})$ and $d' \in (\leq 1\ r)$ as required. Otherwise, assume that such a d'' exists. Since $\mathsf{Tra}(r) \in \mathcal{T}$, and since $\mathcal{I}$ is a model of $\mathcal{T}$ $r^{\mathcal{I}}$ is transitive and thus, $(d, d'') \in r^{\mathcal{I}}$. Then, $d \in (\leq 1\ r)^{\mathcal{I}}$ implies $d' = d''$, from which we can conclude that $d' \in (\exists r.\mathsf{self})^{\mathcal{I}}$. $(d', d'') \in r^{\mathcal{I}}$, and therefore $d' \in (\leq 1\ r)^{\mathcal{I}}$.

For the other direction, let $\mathcal{I}$ be a model of A and $\mathcal{T}'$. Then, define an interpretation $\mathcal{J}$ as follows:

- $\Delta^{\mathcal{J}} := \Delta^{\mathcal{I}}$;
- $o^{\mathcal{J}} := o^{\mathcal{I}}$
- $A^{\mathcal{J}} := A^{\mathcal{I}}$;

- for all simple roles r in $\mathcal{T}$, $r^{\mathcal{J}} := r^{\mathcal{I}}$;
- for non-simple roles r in $\mathcal{T}$,

$$r^{\mathcal{J}} := r^{\mathcal{I}} \cup \bigcup_{s \sqsubseteq^* r, \mathsf{Tra}(s) \in \mathcal{T}} (s^{\mathcal{I}})^+,$$

where $(s^{\mathcal{I}})^+$ denotes the the transitive closure of $s^{\mathcal{I}}$

Claim. For every concept $C \in \mathsf{cl}(\mathcal{T})$, $C^{\mathcal{I}} = C^{\mathcal{J}}$. We show that this claim holds by structural induction on C.

- $C = \{o\}$. Holds trivially by the definition of $\mathcal{J}$.
- $C = A$. The claim follows from the definition of $\mathcal{J}$.
- $C = \neg D$. We have $d \in (\neg D)^{\mathcal{I}}$ iff $d \in \Delta^{\mathcal{I}} \setminus D^{\mathcal{I}}$ iff (by induction hypothesis and definition of $\mathcal{J}$) $d \in \Delta^{\mathcal{J}} \setminus D^{\mathcal{J}}$.
- $C = D_1 \sqcap D_2$ or $C = D_1 \sqcup D_2$. Clearly, if $D_1^{\mathcal{I}} = D_1^{\mathcal{J}}$ and $D_2^{\mathcal{I}} = D_2^{\mathcal{J}}$ by induction hypothesis, we can conclude $(D_1 \sqcap D_2)^{\mathcal{I}} = (D_1 \sqcap D_2)^{\mathcal{J}}$ as well as $(D_1 \sqcup D_2)^{\mathcal{I}} = (D_1 \sqcup D_2)^{\mathcal{J}}$.
- $C = \exists r.D$. The definition of $\mathcal{J}$ ensures that $r^{\mathcal{I}} \subseteq r^{\mathcal{J}}$, for every role r. Then by induction hypothesis $D^{\mathcal{I}} = D^{\mathcal{J}}$, and thus $(\exists r.D)^{\mathcal{I}} \subseteq (\exists r.D)^{\mathcal{J}}$. For the converse, we distinguish two cases: r is a simple role, then $r^{\mathcal{I}} = r^{\mathcal{J}}$, then the induction hypothesis allows us to conclude $d \in (\exists r.D)^{\mathcal{J}}$ implies $d \in (\exists r.D)^{\mathcal{I}}$. On the other hand, if r is a non-simple role, we show that $d \notin (\exists r.D)^{\mathcal{I}}$ implies $d \notin (\exists r.D)^{\mathcal{J}}$. Assume towards a contradiction that $d \in (\exists r.D)^{\mathcal{J}}$, that is, there is some $e \in \Delta^{\mathcal{J}}$ such that $(d, e) \in r^{\mathcal{J}}$ and $e \in D^{\mathcal{J}}$. By assumption $d \notin (\exists r.D)^{\mathcal{I}}$, then either $(d, e) \notin r^{\mathcal{I}}$ or $e \notin D^{\mathcal{I}}$. By induction hypothesis, we know that $e \in D^{\mathcal{I}}$. Then, it must be the case that $(d, e) \notin r^{\mathcal{I}}$. By definition of $\mathcal{J}$ this means that there is a sequence of domain element $d = d_0, \ldots, d_n = e \in \Delta^{\mathcal{I}}$ and a role $s \sqsubseteq^* r$ such that $\mathsf{Tra}(s) \in \mathcal{T}$ and $(d_i, d_{i+1}) \in s^{\mathcal{I}}$ $(0 \leq i < n)$. Observe that since $\exists r.D \in \mathsf{cl}(\mathcal{T})$, then also $D, \neg D \in \mathsf{cl}(\mathcal{T})$ Therefore, $\mathcal{T}'$ contains the axiom

$$\forall s.(\neg D) \sqsubseteq \forall s.(\forall s.(\neg D)).$$

Since $d \notin (\exists r.D)^{\mathcal{I}}$ then $d \in (\forall r. \neq D)^{\mathcal{I}}$. Using a simple inductive argument and because $\mathcal{I} \models \mathcal{T}$ and $s \sqsubseteq^* r$, we can conclude $d_n \in (\neg D)^{\mathcal{I}}$. That is $e \in (\neg D^{\mathcal{I}})$, a contradiction.
- case $C = (\leq 1\ r)$. If r is a simple role, the claim follows by the definition of $\mathcal{J}$, as in this case $r^{\mathcal{I}} = r^{\mathcal{J}}$. Now, let us consider the case r is a non-simple role. Since $r^{\mathcal{I}} \subseteq r^{\mathcal{J}}$, it follows that $(\leq 1\ r)^{\mathcal{J}} \subseteq (\leq 1\ r)^{\mathcal{I}}$. We show next that $(\leq 1\ r)^{\mathcal{I}} \subseteq (\leq 1\ r)^{\mathcal{J}}$. Let $d \in (\leq 1\ r)^{\mathcal{I}}$. Assume towards a contradiction that there are two distinct domain elements $d', d'' \in \Delta^{\mathcal{J}}$ such that $(d, d'), (d, d'') \in r^{\mathcal{J}}$. The definition of $\mathcal{J}$, yields the existence of a role $s \sqsubseteq^* r$ such that $\mathsf{Tra}(s) \in \mathcal{T}$; and the existence of a sequence of domain elements $e_1, \ldots, e_n \in \Delta^{\mathcal{I}}$, with $d = e_1$, $d' = e_2$, $d'' = e_n$, and $(e_i, e_{i+1}) \in s^{\mathcal{I}}$. Since $\mathsf{Tra}(s) \in \mathcal{T}$, $\mathcal{T}'$ contains the axiom

$$(\leq 1\ s) \sqsubseteq \forall s.(\forall s.\bot \sqcup \exists s.\mathsf{self}).$$

As $\mathcal{I}$ is a model of $\mathcal{T}'$ and $s \sqsubseteq^* r$, then $d = e_1 \in (\leq 1\ s)^{\mathcal{I}}$.

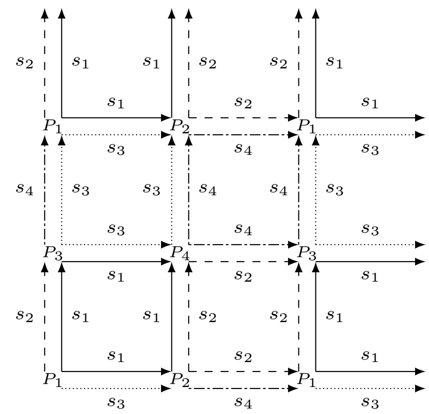

Figure 3: Grid-like model

We can thus conclude that $d \in (\forall s.(\forall s.\bot \sqcup \exists s.\mathsf{self}))^{\mathcal{I}}$, and $d' = e_2 \in (\forall s.\bot \sqcup \exists s.\mathsf{self})^{\mathcal{I}}$. This means that $e_j = e_{j+1}$ for $2 \leq j$, and hence $d' = d''$, a contradiction. We have therefore shown that $d \in (\leq 1\ r)^{\mathcal{J}}$. This finishes the proof of the claim.

We can now show that $\mathcal{J}$ satisfies every CI in $\mathcal{T}$. Indeed, consider an arbitrary axiom $C \sqsubseteq D \in \mathcal{T}$ then $(\neg C \sqcup D)^{\mathcal{I}} = \Delta^{\mathcal{I}}$ as $\mathcal{I}$ is a model of $\mathcal{T}$. Moreover, by the above claim, $(\neg C \sqcup D)^{\mathcal{I}} \subseteq (\neg C \sqcup D)^{\mathcal{J}}$ and hence $(\neg C \sqcup D)^{\mathcal{J}} = \Delta^{\mathcal{J}}$, wich means that $C \sqsubseteq D$ is satisfied in $\mathcal{J}$. Further, every role inclusion axiom $r \sqsubseteq s$ is also satisfied by $\mathcal{J}$: if both r and s are simple roles the claim is a consequence of the definition of $\mathcal{J}$ as $r^{\mathcal{I}} = r^{\mathcal{J}}$ for every role; on the other hand, if both r and s are non-simple, the claim follows from the fact that transitive closure is monotone over set inclusion and because $r^{\mathcal{I}} \subseteq r^{\mathcal{J}}$ for every non-simple role r. Finally, the definition of $\mathcal{J}$ ensures that every transitive role assertion is also satisfied by $\mathcal{J}$. We conclude the proof by noting that, by definition of $\mathcal{J}$, we have $A^{\mathcal{J}} = A^{\mathcal{I}} \neq \emptyset$. ☐

Proofs for Section 6

Theorem 5. *Concept satisfiability relative to DL-Lite$_{core}^{\mathcal{SHN}}$ is undecidable.*

Proof. Given a Turing machine $\mathcal{M}$, we construct a TBox $\mathcal{T}$ such that $\mathcal{M}$ does not accept an input $\vec{w}$ iff a concept $\mathsf{Init}_{\vec{w}}$ is satisfiable relative to $\mathcal{T}$. We obtain the desired undecidability result by applying this construction to a fixed deterministic *universal* Turing machine, i.e., a machine that accepts its input $\vec{w}$ iff the Turing machine encoded by $\vec{w}$ accepts the empty input.

Let $\mathcal{M} = (\Gamma, Q, q_0, q_1, \delta)$ be a deterministic Turing machine, where $\Gamma = \{1, \llcorner\!\lrcorner\}$ is a *two-symbol* tape alphabet, Q is a set of *states*, $q_0 \in Q$ and $q_1 \in Q$ are the *initial* and *accepting* state, respectively, and $\delta \colon Q \times \Gamma \to Q \times \Gamma \times \{-1, +1\}$ is the *transition function*. As usual, computations of $\mathcal{M}$ are sequences of configurations, with each configuration determined by the contents of all (infinitely many) cells of the tape, the state and the head position.

We start by introducing the inclusions of $\mathcal{T}$ describing the basic elements to enforce the infinite grid-like structure. To this aim, we use the following signature

- Concept names $\mathsf{P}_1, \mathsf{P}_2, \mathsf{P}_3, \mathsf{P}_4$ to mark each point on a square on the grid;
- Role names tape_i and next_i, $i \in \{1, \ldots, 4\}$ and tape, next to connect the above points;
- Transitive roles s_i, $i \in \{1, \ldots, 4\}$.

We now have the ingredients to define the inclusions of $\mathcal{T}$ enforcing a grid-like model. We start (eq.(6)) by declaring roles s_i as transitive, enforcing that each point in the grid has at most 3 outgoing s_i roles and ensuring that each point is uniquely marked with a concept P_i.

$$\begin{array}{ll}\mathsf{Tra}(s_i), & (\geqslant 4\ s_i) \sqsubseteq \bot, \\ \mathsf{P}_i \sqcap \mathsf{P}_j \sqsubseteq \bot, & \text{for } i \neq j,\ i, j \in \{1, \ldots, 4\}\end{array} \tag{6}$$

We next relate the roles tape_i and next_i, connecting points in the grid, with the transitive roles via the following RIs. We particularly describe this relation in the `SW, SE, NW` and `NE` quadrant, respectively.

$$\begin{array}{llll}\mathsf{tape}_1 \sqsubseteq s_1, & \mathsf{tape}_1 \sqsubseteq s_3, & \mathsf{tape}_2 \sqsubseteq s_2, & \mathsf{tape}_2 \sqsubseteq s_4, \\ \mathsf{tape}_3 \sqsubseteq s_3, & \mathsf{tape}_3 \sqsubseteq s_1, & \mathsf{tape}_4 \sqsubseteq s_2, & \mathsf{tape}_4 \sqsubseteq s_4, \\ \mathsf{next}_1 \sqsubseteq s_1, & \mathsf{next}_1 \sqsubseteq s_2, & \mathsf{next}_2 \sqsubseteq s_3, & \mathsf{next}_2 \sqsubseteq s_4, \\ \mathsf{next}_3 \sqsubseteq s_1, & \mathsf{next}_3 \sqsubseteq s_2, & \mathsf{next}_4 \sqsubseteq s_3, & \mathsf{next}_4 \sqsubseteq s_4.\end{array}$$

From here on, for a role r, we use $\exists r$ as an abbreviation of $(\geq 1\ r)$.

We are now in the position to construct the grid (see Figure 3 above), properly marking the points and homogeneously naming the relations connecting them (eq. (7)-(9)). We particularly ensure functionality of next, tape and tape^- (eq. (10)), which later is needed to properly synchronise configurations.

$$\begin{array}{ll}\mathsf{P}_1 \sqsubseteq \exists \mathsf{tape}_1 \sqcap \exists \mathsf{next}_1, & \mathsf{P}_2 \sqsubseteq \exists \mathsf{tape}_2 \sqcap \exists \mathsf{next}_3, \\ \mathsf{P}_3 \sqsubseteq \exists \mathsf{tape}_3 \sqcap \exists \mathsf{next}_2, & \mathsf{P}_4 \sqsubseteq \exists \mathsf{tape}_4 \sqcap \exists \mathsf{next}_4\end{array} \tag{7}$$

$$\begin{array}{lll}\exists \mathsf{tape}_1^- \sqsubseteq \mathsf{P}_2, & \exists \mathsf{next}_1^- \sqsubseteq \mathsf{P}_3, & \exists \mathsf{tape}_2^- \sqsubseteq \mathsf{P}_1, \\ \exists \mathsf{next}_3^- \sqsubseteq \mathsf{P}_4 & \exists \mathsf{tape}_3^- \sqsubseteq \mathsf{P}_4, & \exists \mathsf{next}_2^- \sqsubseteq \mathsf{P}_1 \\ \exists \mathsf{tape}_4^- \sqsubseteq \mathsf{P}_3, & \exists \mathsf{next}_4^- \sqsubseteq \mathsf{P}_2 & \end{array} \tag{8}$$

$$\mathsf{tape}_i \sqsubseteq \mathsf{tape} \qquad \mathsf{next}_i \sqsubseteq \mathsf{next} \tag{9}$$

$$\geq 2\, \mathsf{next} \sqsubseteq \bot, \qquad \geq 2\, \mathsf{tape} \sqsubseteq \bot, \qquad \geq 2\, \mathsf{tape}^- \sqsubseteq \bot \tag{10}$$

We now proceed to synchronise successor configurations. We use the following signature:

- Concept names H_q, for each $q \in Q$, to indicate the position of the head of $\mathcal{M}$;
- Concept name $H_\emptyset$ to indicate that the head is not positioned in a cell;
- Role names $\mathsf{next}_{\llcorner\!\lrcorner}$ and next_1, for the two symbols of the alphabet Γ, to represent the content of a cell;
- Role names $\mathsf{next}_{q,-1}$ and $\mathsf{next}_{q,+1}$ to propagate the new state to the next configuration;
- Role names $\mathsf{tape}_{q,+1}$ and $\mathsf{tape}_{q,-1}$ to propagate the head in the state q along the tape;
- Auxiliary role name tape_0 to propagate the blanks beyond the input.

Intuitively, we use the range of next_a, $a \in \Gamma$, to represent a cell that contains a.

A transition $\delta(q, a) = (q', a', \sigma)$ of $\mathcal{M}$ can be naturally encoded using a CI of the form $H_q \sqcap \exists \mathsf{next}_a^- \sqsubseteq \exists \mathsf{next}_{a'} \sqcap \exists \mathsf{next}_{q'\sigma}$ with $\mathsf{next}_{q'\sigma}$ is also a sub-role of next. Note that, however, *DL-Lite*$^{\mathcal{SHN}}_{core}$ does not allow to express Horn-like axioms, that is, we cannot write conjunctions on the left-hand side of CIs. Nevertheless, it has been recently shown in (Gutiérrez-Basulto et al. 2015)[Theorem 13] that using functionality of roles and RIs, available in *DL-Lite*$^{\mathcal{SHN}}_{core}$, one can simulate the above CI in the core fragment. Since the encoding would be virtually the same we do not repeat it here, and instead codify transitions using the mentioned Horn-like CI.

We implement the changes imposed by the transition (eq.(11) and eq.(12)), and ensure that the cells that are not under the head do not change their contents (eq.(13)).

- For $\delta(q, a) = (q', a', \sigma)$,
$$H_q \sqcap \exists \mathsf{next}_a^- \sqsubseteq \exists \mathsf{next}_{a'} \sqcap \exists \mathsf{next}_{q'\sigma} \tag{11}$$
- For $q \in Q$ and $\sigma \in \{-1, +1\}$
$$\mathsf{next}_{q\sigma} \sqsubseteq \mathsf{next} \tag{12}$$
- For $a \in \Gamma$
$$H_\emptyset \sqcap \exists \mathsf{next}_a^- \sqsubseteq \exists \mathsf{next}_a \tag{13}$$

We next place the state variable in the correct position in the successor configuration, where $\mathsf{tape}_{q,\sigma}$ are fresh role names.

$$\exists \mathsf{next}_{q\sigma}^- \sqsubseteq \exists \mathsf{tape}_{q\sigma}, \qquad \text{for } q \in Q \text{ and } \sigma \in \{-1, +1\}, \tag{14}$$

$$\exists \mathsf{tape}_{q\sigma}^- \sqsubseteq H_q, \qquad \text{for } q \in Q \text{ and } \sigma \in \{-1, +1\}, \tag{15}$$

$$\mathsf{tape}_{q,+1} \sqsubseteq \mathsf{tape} \text{ and } \mathsf{tape}_{q,-1} \sqsubseteq \mathsf{tape}^-, \qquad \text{for } q \in Q. \tag{16}$$

Finally, we propagate the no-head marker $H_\emptyset$ (eq.(17)-eq.(19)).

$$H_q \sqsubseteq \exists \mathsf{tape}_{\emptyset,+1} \sqcap \exists \mathsf{tape}_{\emptyset,-1}, \qquad \text{for } q \in Q, \tag{17}$$

$$\mathsf{tape}_{\emptyset,+1} \sqsubseteq \mathsf{tape} \text{ and } \mathsf{tape}_{\emptyset,-1} \sqsubseteq \mathsf{tape}^-, \tag{18}$$

$$\exists \mathsf{tape}_{\emptyset,\sigma}^- \sqsubseteq \exists \mathsf{tape}_{\emptyset,\sigma} \sqcap H_\emptyset, \qquad \text{for } \sigma \in \{-1, +1\}. \tag{19}$$

Next, the following CIs encode an input $\vec{w} = a_1, \ldots, a_n \in \Gamma^*$ of $\mathcal{M}$:

$$\mathsf{Init}_{\vec{w}} \sqsubseteq \mathsf{P}_1 \sqcap C_1 \sqcap \exists \mathsf{headIni} \quad (20)$$

$$C_i \sqsubseteq \exists r_{i+1} \quad \text{for } 1 \leq i < n \quad (21)$$

$$r_i \sqsubseteq \mathsf{tape} \quad \text{for } 1 \leq i \leq n \quad (22)$$

$$\exists r_i^- \sqsubseteq C_i \quad \text{for } 1 < i \leq n \quad (23)$$

$$C_i \sqsubseteq \exists \mathsf{next}_{a_i} \quad \text{for } 1 \leq i \leq n \quad (24)$$

$$\exists \mathsf{headIni} \sqsubseteq H_{q_0} \quad (25)$$

We fill the rest of the tape by blanks:

$$C_n \sqsubseteq \mathsf{tape}_0 \quad (26)$$

$$\exists \mathsf{tape}_0^- \sqsubseteq \exists \mathsf{next}_{\llcorner} \sqcap \exists \mathsf{tape}_0, \qquad \mathsf{tape}_0 \sqsubseteq \mathsf{tape}. \quad (27)$$

We ensure that the accepting state q_1 never occurs in a computation:

$$H_{q_1} \sqsubseteq \bot. \quad (28)$$

This finishes the construction of $\mathcal{T}$. Following the intuitions above it is not to hard see that $\mathsf{Init}_{\vec{w}}$ is satisfiable relative to $\mathcal{T}$ iff $\mathcal{M}$ does not accept the word $\vec{w}$. □

In the results below we denote with $\mathcal{F}$ the presence of *(global) functionality axioms* of the form $\mathsf{func}(r)$, with the usual semantics: $\mathcal{I} \models \mathsf{func}(r)$ iff $e_1 = e_2$ for all $(d, e_1), (d, e_2) \in r^{\mathcal{I}}$. Recall that we denote with *DL-Lite*$_{bool}^{\mathcal{F},\mathsf{sf}}$ the extension of *DL-Lite*$_{bool}^{\mathcal{F}}$ with local reflexivity concepts (cf. Section 5)

Lemma 4 *Concept satisfiability relative to DL-Lite*$_{bool}^{\mathcal{F},\mathsf{sf}}$ *TBoxes is* NP*-complete.*

Proof. The lower bound is inherited from *DL-Lite*$_{bool}^{\mathcal{F}}$. For the upper bound, one can extend the reduction from satisfiability in *DL-Lite*$_{bool}^{\mathcal{F}}$ to satisfiability in the one-variable fragment of first-order logic ($\mathcal{QL}^1$) (Artale et al. 2009, Section 5.1), in order to deal with local reflexivity. □

Theorem 7 *Concept satisfiability relative to DL-Lite*$_{bool}^{\mathcal{S},\mathcal{F}}$ *TBoxes is* NP*-complete.*

Proof. The lower bound is inherited from *DL-Lite*$_{bool}^{\mathcal{F}}$. For the upper bound, we polynomially reduce satisfiability in *DL-Lite*$_{bool}^{\mathcal{S},\mathcal{F}}$ to satisfiability in *DL-Lite*$_{bool}^{\mathcal{F},\mathsf{sf}}$ as follows:

Let $\mathcal{T}$ be a *DL-Lite*$_{bool}^{\mathcal{S},\mathcal{F}}$ TBox. We obtain a *DL-Lite*$_{bool}^{\mathcal{F},\mathsf{sf}}$ TBox $\mathcal{T}'$ from $\mathcal{T}$ by taking $\mathcal{T}' = (\mathcal{T} \setminus \{\mathsf{Tra}(r) \in \mathcal{T}\}) \cup \mathcal{T}''$, where $\mathcal{T}''$ is the set of the following CIs, for every r such that $\{\mathsf{Tra}(r), \mathsf{func}(r)\} \subseteq \mathcal{T}$:

$$\exists r^- \sqsubseteq \neg \exists r \sqcup \exists r.\mathsf{self}$$

One can show the correctness of this reduction similar to Lemma 7 above:

Lemma 8. *For every concept name A, A is satisfiable relative to $\mathcal{T}$ iff A is satisfiable relative to $\mathcal{T}'$.*

□